4.0   TRENBOLONE

4.1 EPA 21-Day Assay for Methoxychlor

The EPA 21-day trenbolone assay was conducted from January 21, 2003 to February 3, 2003 (prevalidation test) and from February 3, 2003 to February 24, 2003 (validation test).

4.1.1   Chemistry

Trenbolone concentrations were measured in each replicate on Day 4, 7, 14, and 21 (Figure 4.1). An additional replicate of the Low concentration was sampled on Day 10. trenbolone concentrations in the Control treatment are reported at the method detection limit (MDL). The mean percent difference of the High and Low concentrations achieved during the test versus the target concentrations were 40% and 59%, respectively (Table 4.1).

Table 4.1. Summary statistics for trenbolone concentrations (μg/L) during the EPA 21-Day assay

	Treatment		
	Control	**Low**	**High**
Target	0.051 [1]	0.1	1
mean	0.051 U	0.041 [2]	0.60
stdev	0	0.012	0.19
min	0.051 U	0.013	0.33
max	0.051 U	0.060	0.92
median	0.051 U	0.042	0.33
Q1	0.051 U	0.033	0.50
Q3	0.051 U	0.051	0.70
Mean % Δ	–	59%	40%

[1] Equal to the MDL.
[2] Concentrations in some replicates were estimated at levels less than the MDL.
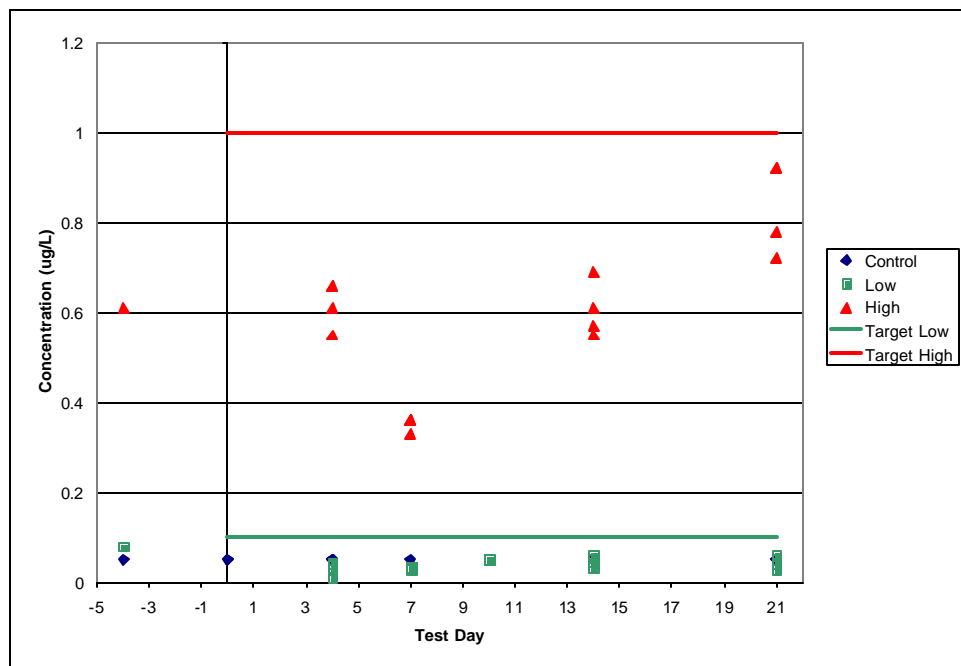U Not detected at concentrations greater than the MDL.

Figure 4.1. trenbolone concentrations measured in the High and Low-concentration treatments versus the target concentrations (1.0 µg/L and 0.1 µg/L) during the EPA 21-day trenbolone assay. The Control treatment values are plotted at the MDL.

4.1.2 Assay Conditions Summary

All water quality, water chemistry, and light intensity parameters were within the acceptable ranges during the EPA 21-day trenbolone prevalidation and validation tests (Table 4.2). Flow rates were monitored regularly, and were adjusted if they were observed to be out of the target range (see Methods).

Table 4.2. Summary of water quality, water chemistry, flow rate, and light intensity data for the EPA 21-day trenbolone assay

	Temperature (°C)		Dissolved Oxygen (mg/L)		pH		Hardness (mg/L)		Alkalinity (mg/L)		Ammonia (mg/L)		Light Intensity (lux)			
	min	max	min	max	min	max	min	max	min	max	min	max	min	max	mean	SD
Pre-exposure Assay																
Target Range	24	26	>4.9		6.5	9.0	>140		>20		<0.5		540	1080		
Control	24.0	25.8	6.5	7.9	7.3	8.3	153	170	150	160	0.0	0.2	760	875	817	46
Low	24.0	25.3	6.8	8.0	7.7	8.3	170	187	150	160	0.0	0.2	708	1039	807	110
High	24.1	24.9	6.7	8.0	7.2	8.3	170	187	140	160	0.0	0.3	713	1009	809	87
Exposure Assay																
Target Range	24	26	>4.9		6.5	9.0	>140		>20		<0.5		585	1060	906	154
Control	23.9	25.2	5.8	7.6	7.3	8.2	170	221	130	160	0.1	0.7	642	1065	927	194
Low	23.9	25.1	6.4	7.7	7.1	8.5	170	204	140	160	0.1	0.8	808	1057	956	120
High	23.9	25.1	6.3	7.6	7.3	8.2	170	204	150	160	0.7	0.7	840	1078	976	116

4.1.3 Survival

All males in all treatments and all females in the Control treatment survived the EPA 14-day Methoxychlor assay. One female in the High concentration and one female in the Low-concentration treatments died during the 14-day exposure (94% survival in each case).

4.1.4 Fecundity

Total Fecundity—Two pre-exposure evaluations of total egg production were performed. One examined the total egg production at 14 days, which is within the protocol guidelines; the other evaluated egg production at day 7 of the pre-exposure assay to determine whether or not the shorter pre-exposure period would be sufficient to evaluate the potential reproductive success of the test fish. Total 14-day counts among the tanks that were eventually used for the three treatments in the exposure assay (individual tank values summed for each treatment) ranged from about 12,000 eggs to 14,000 eggs (Figure 4.2). No significant differences in the mean 14-day egg production among the groups of replicates that would be used in the trenbolone exposure assay were detected (1-way ANOVA, $F = 0.11$, $p = 0.953$, df = 3, 13). Production at Day 7 of the pre-exposure test among the 12 replicates that were eventually used in the exposure assay ranged from 503 eggs to 1,923 eggs. No significant differences in the mean 7-day egg production among the groups of replicates that would be used in the trenbolone exposure assay were detected (1-way ANOVA, $F = 0.11$, $p = 0.955$, df = 3, 13). No significant differences were detected in the mean number of eggs laid per day per replicate group at 7 days versus 14 days (Two-Sample *t*-Test, $t = -0.85$, $p = 0.401$, df = 25). However, daily within-treatment variation (indicated by fluctuating coefficient of variation values) was high through Day 9. Therefore, variability during a 7-day pre-exposure assay is likely to be much greater than that during longer pre-exposure assay, which reduces the likelihood of reliably choosing replicates with successful reproduction histories for use in the exposure assay.
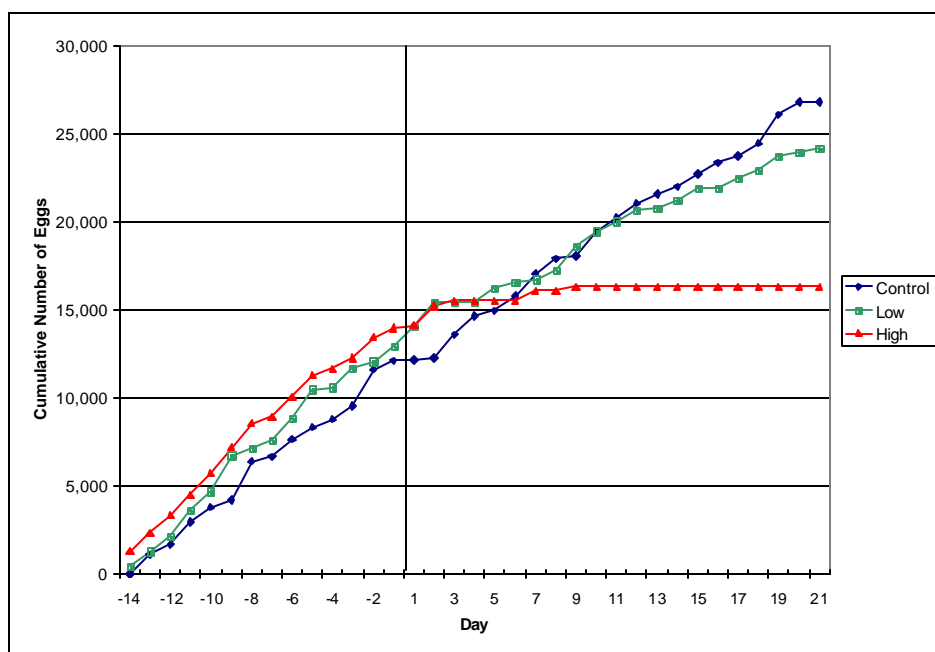
Figure 4.2. Total egg production per treatment for the EPA 21-Day trenbolone assay. The vertical line at Day 0 denotes the start of the exposure period.

During the 21-day trenbolone assay, total counts in the Control treatment were reasonably consistent among replicates, varying from 2,741 eggs to 4,334 eggs (**Figure E21T Fec-2**). Variability in total egg production among Low concentration replicates was somewhat greater, ranging from 1,873 eggs to 4,017 eggs. Total counts among the High concentration replicates varied about five-fold, ranging from 215 eggs to 1,095 eggs. Statistical analysis of square-root transformed egg counts showed significant among-treatment differences (1-way ANOVA, $F = 23.14$, $p < 0.001$, df = 2, 10) in mean cumulative numbers of eggs produced (**Table E21T Fec-1**). Tukey's pairwise comparison identified significant differences in mean egg production between the High trenbolone concentration and the Control treatment, and also between the High and Low concentrations. The Low concentration and the Control treatment fecundity values were not statistically different. Daily within-treatment variation (indicated by fluctuating coefficient of variation values) was much reduced after Days 8–10, except for the High concentration. High mortality occurred in one replicates of the High trenbolone concentration. All fish in this replicate expired by Day 7 of the exposure. Egg production in the other four High concentration replicates was sporadic. Egg production stopped after Day 3 in two replicates, after Day 7 in a third replicate, and occurred only on Day 9 in the fourth replicate. The achieved power for this assay was 100%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 3 (**Table E21T Fec-1**).

Table E21T Fec-1. Summary statistics, detectable difference, and power estimates for fecundity data for the EPA 21-day trenbolone assay

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
Control	4	3672	795	22%	100%	3
Low	4	2818	914	32%		
High	4	590	426.5	72%		

[1] Calculated from square-root transformed data; with sample size = 4.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on square-root transformed data.
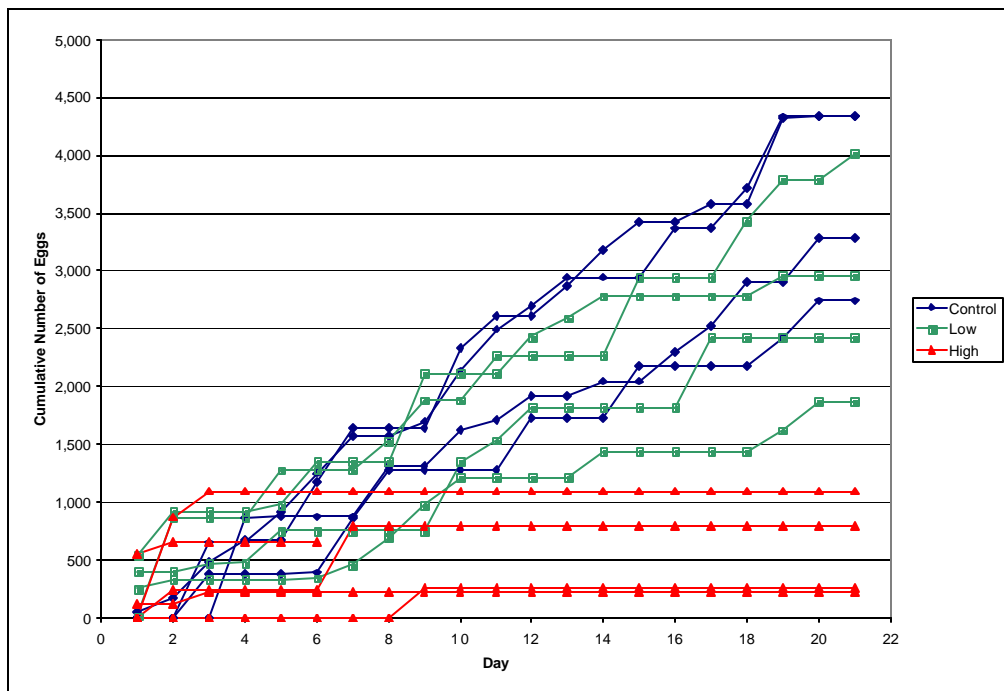

Figure E21T Fec-2. Total egg production by replicate per treatment for the EPA 21-Day trenbolone assay.

Fecundity per Female Reproductive Day—During the 14-Day pre-exposure evaluation, the mean number of eggs produced per female reproductive day ranged from 55.6 eggs/day for the tanks that would be used for the Control treatment to 58.5 eggs/day for the tanks that would be used for the High concentration during the 21-day exposure assay. There were no significant differences among treatments in the numbers of eggs produced per reproductive day during the pre-exposure period (Kruskal-Wallis, $H = 0.99$, $p = 0.804$, df = 3).

During the EPA 21-day trenbolone assay, the maximum number of female reproductive days was achieved for the Control treatment, whereas 80 and 79.3 female reproductive days were achieved in the Low concentration and High concentration, respectively (**Table E21T Fec-2**). The number of eggs produced per female reproductive day varied from 32.6 eggs to 51.6 eggs in the Control

treatment and from 27.5 to 47.8 in the Low concentration (**Figure E21T FEC-2**). For the High concentration, the number of eggs produced per female reproductive day ranged from 2.6 eggs to 13.0 eggs. A significant difference in the mean number of eggs produced per female reproductive day among treatments was detected (Kruskal-Wallis, $H = 8.38$, $p = 0.015$, df = 2). The mean number of eggs produced per female reproductive day was significantly smaller in the High concentration than in either the Control treatment or the Low concentration. The mean number of eggs produced in the Low concentration did not differ from that produced in the Control treatment. The achieved power for this assay was 100%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 3 (**Table E21T Fec-2**).

Table E21T Fec-2. Summary statistics, detectable difference, and power estimates for fecundity per female reproductive day for the EPA 21-day trenbolone assay

Level	Mean Number of Reproductive Days [1]	N	Mean	Stdev	CV	Achieved Power [2]	Sample Size Required [3]
control	84.0	4	43.7	9.5	22%	100%	3
low	80.0	4	34.8	9.3	27%		
high	79.3	4	7.2	4.8	67%		

[1] Maximum number = 84.
[2] Calculated from natural log transformed data; with sample size = 4.
[3] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
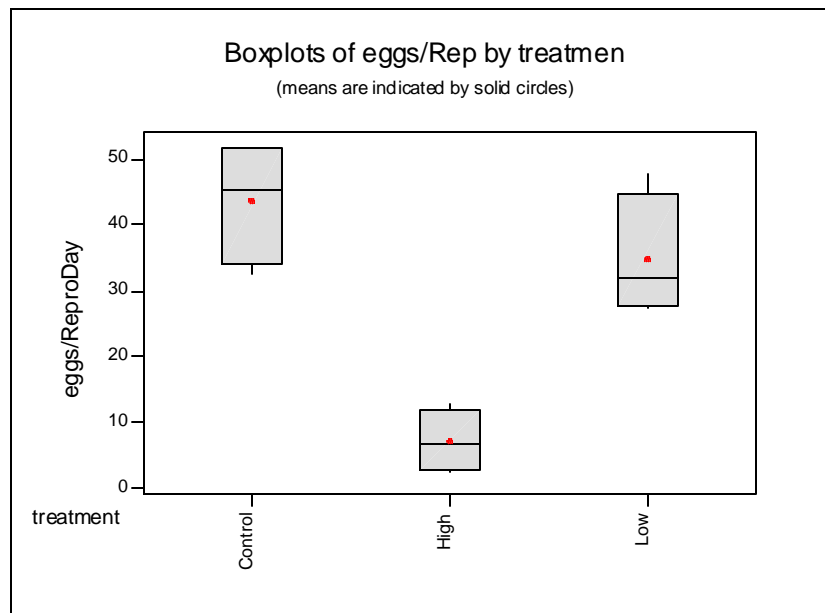

Figure E21T FEC-2. Boxplot of the number of eggs produced per female reproductive day by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Eggs on Tiles/Dishes—The mean number of eggs laid on the tiles during the 14-day pre-exposure assay varied from 2,101 eggs for the tanks that would be used for the Control treatment to 2,561 eggs for the tanks that would be used for the High-concentration treatment. The mean number of eggs laid on the dishes during the 14-day pre-exposure assay varied from 733 eggs for the tanks that would be used for the High-concentration treatment to 1,005 eggs for the tanks that would be used for the Control treatment. Because of the variability in the total number of eggs laid per treatment, the proportional difference in the number of eggs on dishes versus those on tiles [1−(# eggs on dishes ÷ # eggs on tiles)] was calculated. There were no significant differences in the mean proportional difference among treatments during the 14-day pre-exposure assay (Kruskal-Wallis, $H = 1.39$, $p = 0.707$, df = 3).

The mean number of eggs laid on the tiles among the treatments during the EPA 21-day trenbolone assay varied from 500 eggs for the High concentration to 3,384 eggs for the Control treatment (**Table E21T FEC-3**). The number of eggs on dishes ranged from 55 eggs for the High concentration to 130 eggs for the Control treatment. The proportional difference ranged from −0.10 (one High concentration replicate had more eggs on the dish than on the tile) to 0.96 (one Control-treatment replicate) (**Figure E21T FEC-3**). There were no significant differences in this mean proportional difference among treatments (Kruskal-Wallis, $H = 2.46$, $p = 0.292$, df = 2).

Table E21T FEC-3. Summary statistics for the numbers of eggs laid on tiles versus those in dishes for the EPA 21-day trenbolone assay

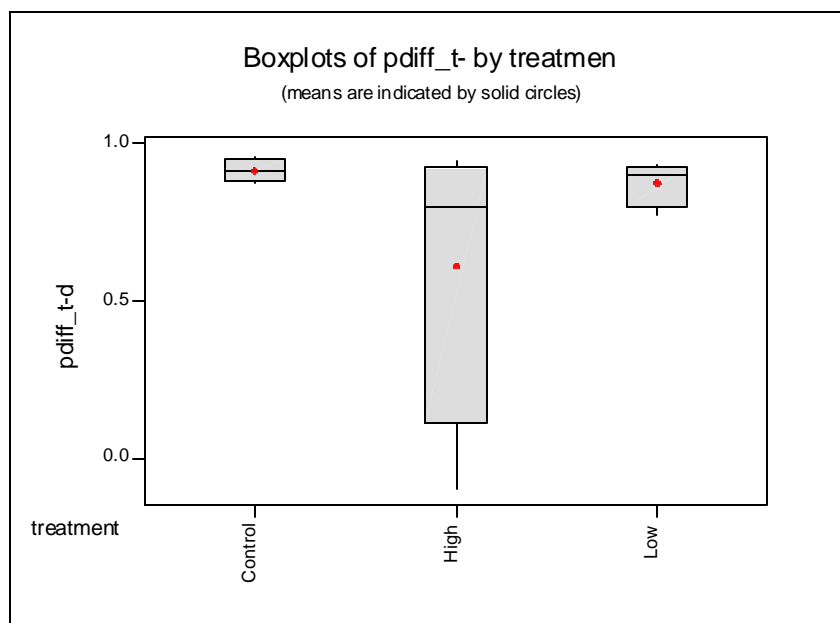|		Tile			Dish			Mean Proportional Difference
Level	N	Mean	Stdev	CV	Mean	Stdev	CV	
control	4	3384	722	21%	288	130	45%	0.92
low	4	2547	941	37%	272	62	23%	0.88
high	4	500	416	83%	90	55	61%	0.61

Figure E21T FEC-3. Boxplot of the proportional difference in the number of eggs laid on tiles versus those in dishes by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

4.1.5 Fertilization Success

Total Fertilization— Eggs were collected during the 14-day pre-exposure period for the evaluation of fertilization-success rate. The mean proportion of eggs fertilized in the Control treatment was 0.981 [standard deviation (sd) = 0.028], 0.997 (sd = 0.002) in the Low concentration, and 0.985 (sd = 0.023) in the High concentration. The mean proportion of eggs fertilized in the replicates that were not used in the 21-day validation assay was 0.977 (sd = 0.041). There were no significant differences among treatments in the proportion of eggs that were fertilized (Kruskal-Wallis, $H = 2.42$, $p = 0.490$, df = 3).

The total (tiles + dishes) fertilization-success rates for all treatment replicates during the EPA 21-day trenbolone assay were high, ranging from 0.991 (one Low-concentration replicate) to 1.00 (one High-concentration replicate) (**Figure E21T FER-1**). No significant differences in mean fertilization-success rates (**Table E21T FER-**1) among treatments were detected (Kruskal-Wallis, $H = 0.04$, $p = 0.981$, df = 2). The achieved power for this assay was 5%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 764 (**Table E21T FER-1**).

Table E21T FER-1. Summary statistics, detectable difference, and power estimates for the proportion of eggs fertilized for the EPA 21-day trenbolone assay

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
Control	4	0.997	0.001	0.1%	5%	764
Low	4	0.996	0.004	0.4%		
High	4	0.997	0.004	0.4%		

[1] Calculated from arcsine square-root transformed data; with sample size = 4.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
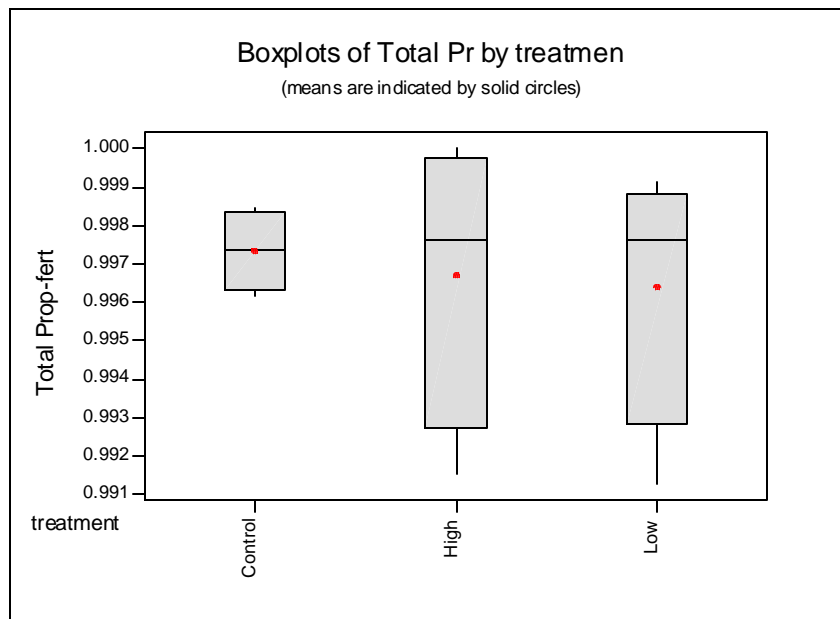

Figure E21T FER-1. Boxplot of the proportion of eggs fertilized by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Fertilization of Eggs on Tiles and Dishes—During the 14-Day pre-validation assay, there were no significant differences in the fertilization-success rates among treatments for eggs laid in tiles (Kruskal-Wallis, $H = 4.51$, $p = 0.211$, df = 3) or on dishes (Kruskal-Wallis, $H = 2.67$, $p = 0.446$, df = 3).

The fertilization-success rates for all treatment replicates for eggs laid on tiles during the EPA 21-day trenbolone assay were high, ranging from 0.990 (one Low-concentration replicate) to 1.00 (two High-concentration replicates) (**Figure E21T FER-2**). No significant differences in mean fertilization-success rates (**Table E21T FER-2**) among treatments were detected (Kruskal-Wallis, $H = 0.35$, $p = 0.841$, df = 2). The fertilization-success rates for all treatment replicates for eggs laid on dishes during the assay were high, ranging from 0.958 (one Control-treatment replicate) to

1.00 (several replicates; including all treatments) (**Figure E21T FER-3**). No significant differences in mean fertilization-success rate s (**Table E21T FER-3**) among treatments were detected (Kruskal-Wallis, $H = 0.62$, $p = 0.733$, df = 2).

Table E21T FER-2. Summary statistics for the proportion of eggs on tiles versus those in dishes that were fertilized for the EPA 21-day trenbolone assay

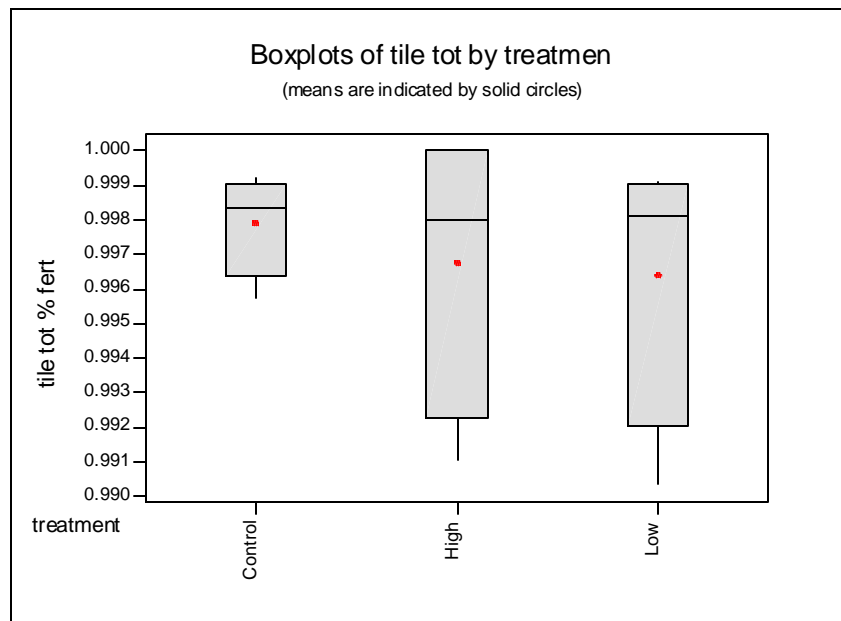| Level	N	Tiles			Dishes		
		Mean	Stdev	CV	Mean	Stdev	CV
control	4	0.998	0.001	0%	0.987	0.020	2%
low	4	0.996	0.004	0%	0.994	0.012	1%
high	4	0.997	0.004	0%	0.996	0.004	0%

Figure E21T FER-2. Boxplot of the proportion of eggs on tiles that were fertilized by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Figure E21T FER-3. Boxplot of the proportion of eggs in dishes that were fertilized by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.
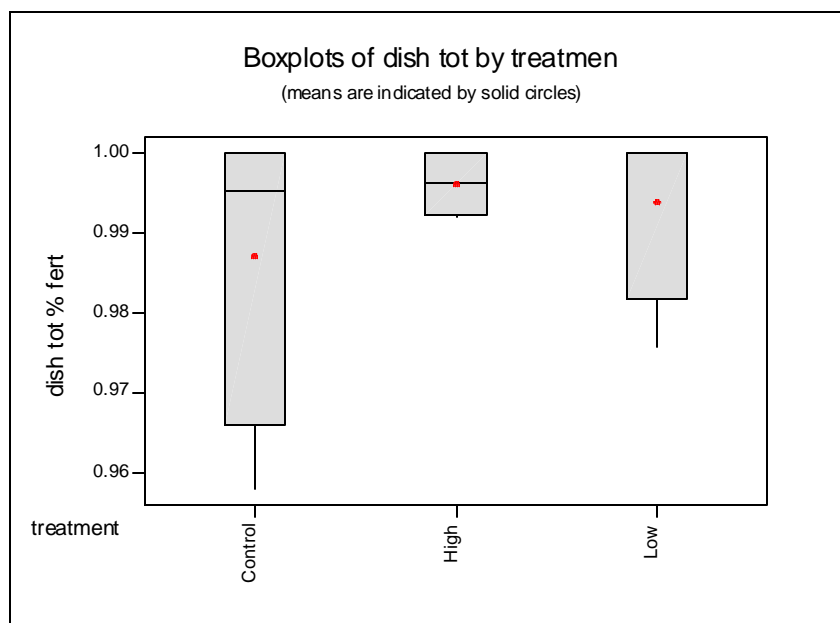
4.1.6   Hatchability and Larval Development

Eggs were collected during the pre-exposure period for the evaluation of hatchability. The mean proportion of fertilized eggs that hatched was 0.97 (sd = 0.02) for Control treatment, 0.98 (sd = 0.02) for the Low concentration, and 0.94 (sd = 0.06) for the High concentration. Among the tanks evaluated during the pre-exposure period, but that were not used in the 21-Day assay, the mean proportion of fertilized eggs that hatched was 0.92 (sd = 0.03). There were no significant differences among treatments in the proportion of eggs that hatched (Kruskal-Wallis, $H = 6.91$, $p = 0.075$, df = 3).

Eggs were collected during the EPA 21-day trenbolone assay for the evaluation of hatchability. The proportion of fertilized eggs that hatched ranged from 0.26 to 1.00 in the Control treatment and from 0.82 to 1.00 for the two test concentrations (**Figure E21T HAT-1**). There were no significant differences among treatments in the proportion of eggs that hatched (Kruskal-Wallis, $H = 0.07$, $p = 0.964$, df = 2). The achieved power for this endpoint was 5%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 185 (**Table E21T HAT-1**). One of the Control treatment replicates had a very low proportion of the eggs hatch (0.26) during the second week of the hatch evaluation. Because the data sheet for that replicate indicated that the eggs may have been abnormal at the initiation of the evaluation, the analyses were rerun excluding those data. The resulting mean proportion hatch for the Control treatment was 0.95 (sd = 0.08), however there still were no significant differences among treatments in the proportion of eggs that hatched (Kruskal-Wallis, $H = 0.59$, $p = 0.746$, df = 2).

Table E21T HAT-1. Summary statistics, detectable difference, and power estimates for the proportion of fertile eggs that hatched for the EPA 21-day trenbolone assay

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	0.86	0.26	30%	5%	185
low	8	0.94	0.06	7%		
high	2	0.94	0.08	9%		

[1] Calculated from arcsine square-root transformed data; with sample size = 2.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
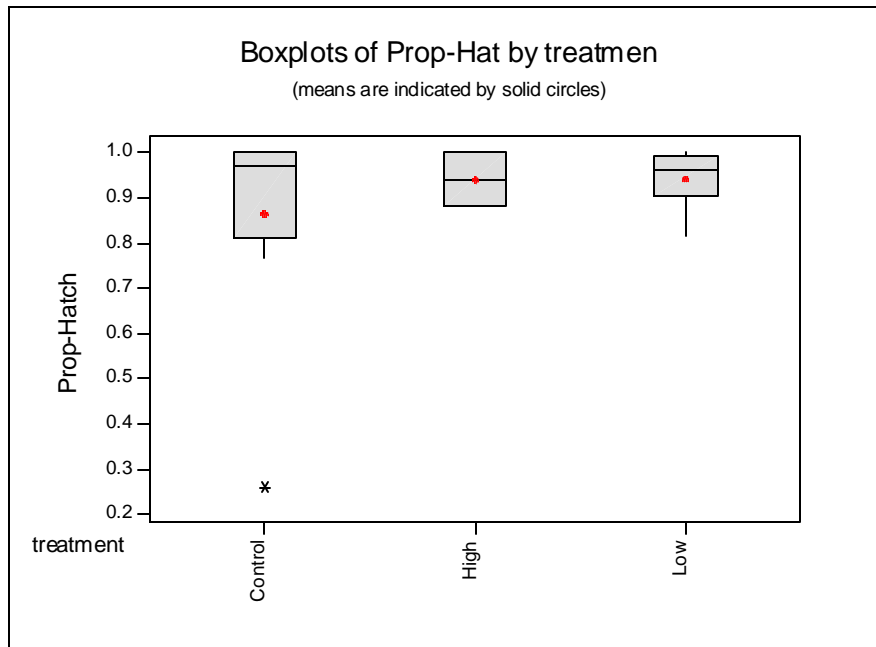

Figure E21T HAT-1. Boxplot of the proportion of fertile eggs that hatched by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, asterisks represent probable outliers.

Eggs were collected during the pre-exposure period for the evaluation of larval development. The proportion of larvae that developed normally (i.e., that showed no morphological abnormalities) was 1.00 for all replicates of the Control treatment. The mean proportion of normal larvae in the remaining treatments was 0.96 (sd = 0.01) in the Low concentration and 0.95 (sd = 0.07) in the High concentration. Among the tanks evaluated during the pre-exposure period, but that were not used in the 21-Day assay, the mean proportion of normal larvae was 0.93 (sd = 0.05). There were no significant differences among treatments in the proportion of normal larvae (Kruskal-Wallis, $H = 6.26$, $p = 0.099$, df = 3).

Eggs were collected on Days 7–10, 14–16, and 21 during the EPA 21-day trenbolone assay for the evaluation of larval development. The proportion of larvae that developed normally (i.e., that showed no morphological abnormalities) ranged from 0.08 to 1.00 in the Control treatment and from 0.98 to 1.00 for the two test concentrations (**Figure E21T HAT-2**). There were no significant differences among treatments in the proportion of normal larvae (Kruskal-Wallis, $H = 2.77$, $p = 0.250$, df = 2). The achieved power for this endpoint was 9%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 16 (**Table E21T HAT-2**). As for the hatchability analyses, the larval development analysis was rerun excluding the one Control treatment replicate that may have had abnormal eggs at the initiation of the evaluation. The resulting mean proportion of larvae that developed normally for the Control treatment was 0.98 (sd = 0.22), however there were no significant differences among treatments in the proportion of normal larvae (Kruskal-Wallis, $H = 1.89$, $p = 0.390$, df = 2).

Table E21T HAT-2. Summary statistics, detectable difference, and power estimates for the proportion of normal larvae for the EPA 21-day trenbolone assay

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	0.78	0.35	45%	9%	16
low	8	0.997	0.01	1%		
high	2	0.99	0.02	2%		

[1] Calculated from arcsine square-root transformed data; with sample size = 2.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
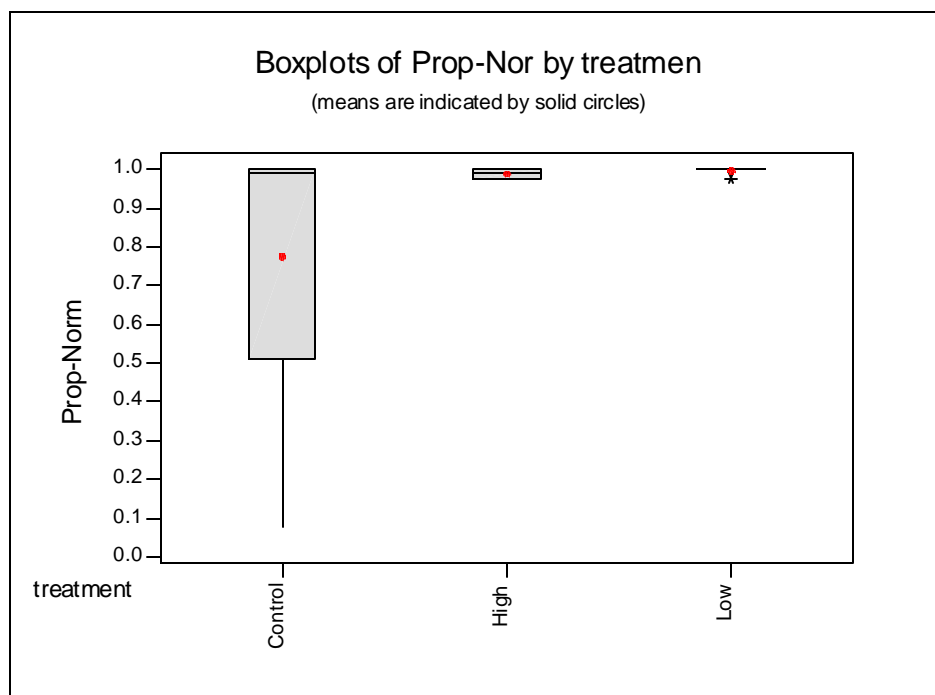
Figure E21T HAT-2. Boxplot of the proportion of normal larvae by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

4.1.7   Body Weight

The body weight of females used in the EPA 21-day trenbolone assay ranged from 1.1 g to 3.7 g. A significant difference in the mean female body weight per treatment was detected (Kruskal-Wallis, $H = 16.34$, $p < 0.001$, df = 2). The test indicated that the mean female body weight for females in the High-concentration treatment was greater than those for the other two treatments. The body weight of males used in the EPA 21-day trenbolone assay ranged from 2.6 g to 6.0 g. There were no significant differences in mean body weight among treatments (Kruskal-Wallis, $H = 0.97$, $p = 0.616$, df = 2).

4.1.8   Gonadosomatic Index

The range of GSI values calculated for females in the all treatments varied from 4- to 22-fold (**Figure E21T GSI-1**), and the overall within-treatment variability, as indicated by CVs, ranged from 34%–52%; **Table E21T GSI-1**). Several fish had relatively high GSI values. The highest value was obtained for a female exposed to the Low trenbolone concentration had a GSI value of 29.8. The GSI values for several other fish exceeded 21.0, which is much greater than the typical upper limit for reproductively-active female fathead minnows (GSI ~13.0). The total body weight of these females were well within the normal range for the test animals and it appears the high GSI value resulted from each fish having unusually high gonad weights. The gonad weights of these fish ranged from 0.41 g to 0.76 g (average = 0.59 g), greater than the average values for the remaining fish (0.2 g). A significant difference in the mean GSI value per treatment (**Table E21T GSI-1**) was detected (Kruskal-Wallis, $H = 8.45$, $p = 0.015$, df = 2). The test indicated that the mean GSI for females in the Control treatment was less than those for the other two

treatments. However, the mean GSI values calculated for the Low and High trenbolone concentrations were not statistically different. Because female body weights were not different between the two treatments (previous section), the gonads of females in the Low concentration were proportionally larger (i.e., hypertrophied) compared to those in the High concentration. The achieved power for this endpoint was 58%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 24 (**Table E21T GSI-1**).

Table E21T GSI-1. Summary statistics, detectable difference, and power estimates for female gonadosomatic index data for the EPA 21-day trenbolone assay

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	16	10.4	5.4	52%	58%	24
low	15	15.9	5.4	34%		
high	15	15.9	7.7	48%		

[1] Calculated from arcsine square-root transformed data; with sample size = 15.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
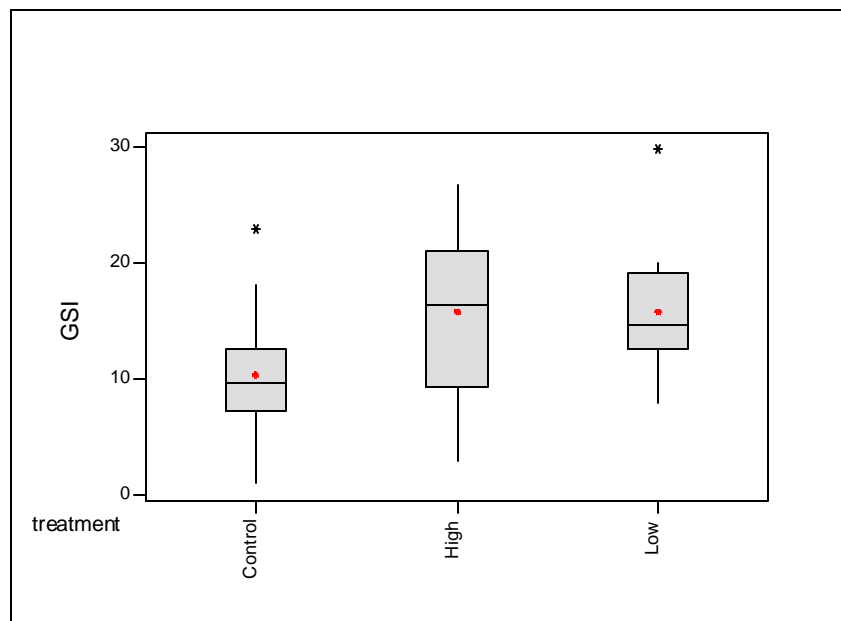

Figure E21T GSI-1. Boxplot of female GSI by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisks represent probable outliers.

The range of most GSI values calculated for males during the EPA 21-day trenbolone assay, was small, ranging from 0.4 to 1.7 (**Figure E21T GSI-2**). Most estimates were between about 1.0 and 1.7, which approximates the typical range for reproductively-active male fathead minnows. There were no significant differences in mean GSI values (**Table E21T GSI-2**) among treatments (Kruskal-Wallis, $H = 0.97$, $p = 0.617$, df = 2). The achieved power for this endpoint was 19%,

and the sample size required to detect a significant difference from the Control treatment at 80% power was 42 (**Table E21T GSI-2**).

Table E21T GSI-2. Summary statistics, detectable difference, and power estimates for male gonadosomatic index data for the EPA 21-day trenbolone assay

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	1.12	0.47	42%	19%	42
low	8	1.20	0.30	25%		
high	8	1.35	0.27	20%		

[1] Calculated from arcsine square-root transformed data; with sample size = 8.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
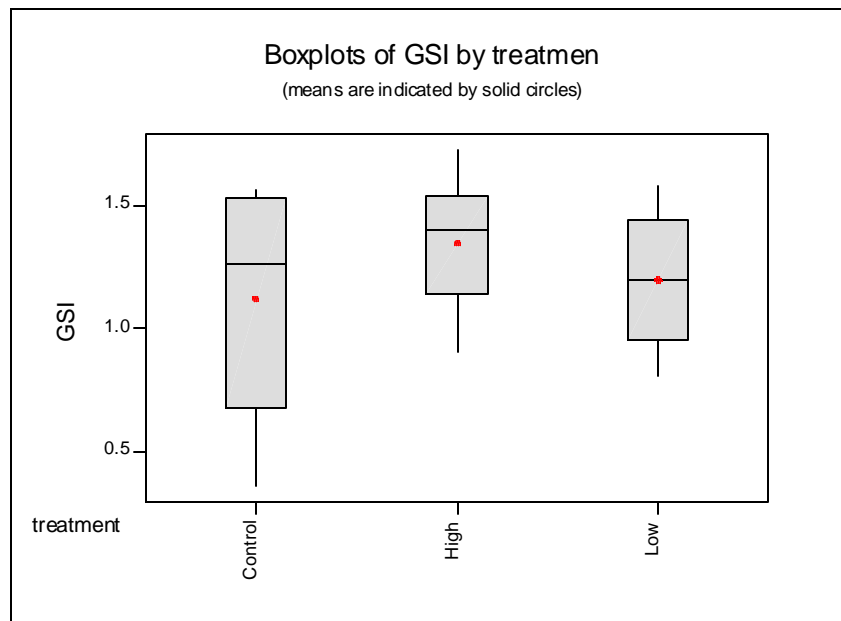

Boxplots of GSI by treatmen
(means are indicated by solid circles)

Figure E21T GSI-2. Boxplot of male GSI by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

4.1.9   Female Gonad Histology

General Ovary Staging—Statistical analysis of the mean ovarian staging from 12 microscopic fields per female in the EPA 21-day trenbolone assay revealed no significant differences among treatments (Kruskal-Wallis, $H = 4.71$, $p = 0.095$, df = 2).

Quantitative Ovarian Staging—One hundred cells in each of three sections per female were examined to quantitatively determine the developmental stage of the ovaries. Ova from fish in all treatments ranged from Stage 1a to Stage 5 (see Methods for a description of the stages) (**Figure E21T FHist-1**). Variability within treatments for each stage was very high as indicated by CVs

that ranged as high as 412% (**Table E21T FHist-1**). Although statistical analyses showed that there was a significant difference among treatments in the proportion of cells in developmental Stages 1a, 1b, 2, and 3, there were no significant differences among treatments in the proportion of cells in the developmental Stages 4, and 5 (**Table E21T FHist-1**). There was no consistent pattern of significant difference associated with trenbolone dose.

Table E21T FHist-1. Descriptive statistics of the proportion of ovarian cells in each developmental stage for females from the EPA 21-day trenbolone assay and results of the Kruskal-Wallis Test (df = 2) comparing treatments

| Stage	Control (n = 16)			Low (n = 13)			High (n = 17)			Kruskal-Wallis	
	Mean	**Stdev**	**CV**	**Mean**	**Stdev**	**CV**	**Value**	**Stdev**	**CV**	*H*	*p*
1a	0.079	0.034	44%	0.099	0.021	22%	0.077	0.019	25%	6.94	0.031*
1b	0.365	0.065	18%	0.327	0.067	20%	0.265	0.089	34%	8.36	0.015*
2	0.202	0.057	28%	0.178	0.029	16%	0.250	0.096	38%	8.03	0.018*
3	0.102	0.052	51%	0.122	0.031	26%	0.192	0.096	50%	14.09	0.001***
4	0.215	0.096	45%	0.240	0.068	28%	0.156	0.106	68%	5.47	0.065
5	0.003	0.008	278%	0.001	0.003	361%	0.002	0.006	412%	2.76	0.252
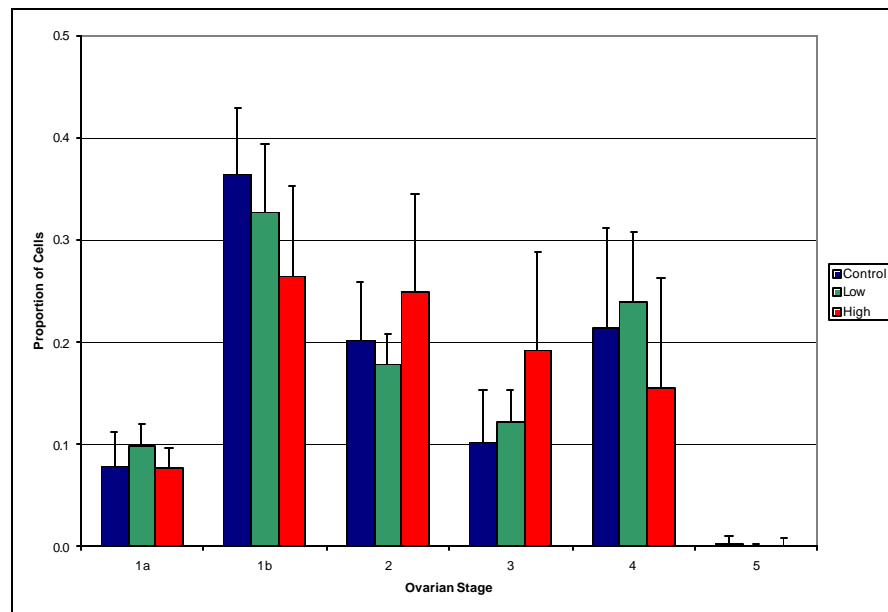
Figure E21M FHist-1. Frequency histogram showing the quantitative developmental staging of ovaries for each treatment of the EPA 21-day trenbolone assay. For each treatment, the columns represent the grand mean proportion of cells in each stage and the bars represent the standard deviation.

Atretic Follicles—The mean proportion of atretic follicles per 300 follicles (counted per fish) ranged from 0.007 for females in the Control treatment to 0.05 for females in the High concentration (**Figure E21T FHist-2**). There were significant differences in the proportions of atretic follicles among treatments (Kruskal-Wallis, $H = 14.88$, $p = 0.001$, df = 2). The proportions of atretic follicles in the High concentration were greater than those in the Control treatment, but not greater than those in the Low-concentration treatment. Therefore, there was no consistent pattern of significant difference associated with trenbolone dose.
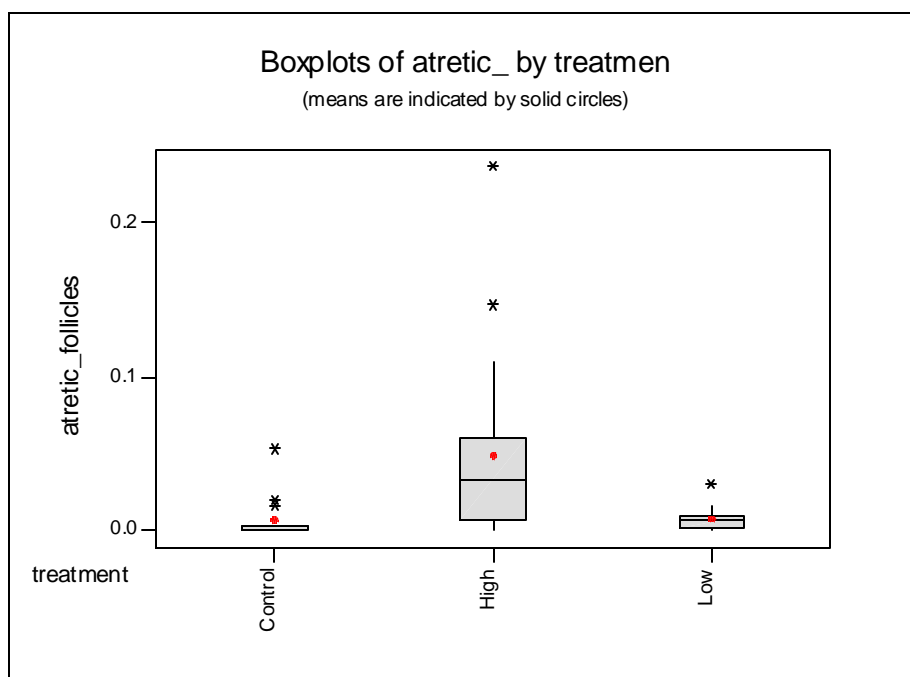
Figure E21T FHist-2. Boxplot of the proportion of atretic follicles per 300 follicles by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and asterisks represent probable outliers.

Corpora Lutea—The mean proportion of corpora lutea per 300 follicles (counted per fish) ranged from 0.01 for females in the High concentration to 0.025 for females in the Low concentration (**Figure E21T FHist-3**). There were significant differences in the proportions of corpora lutea among treatments (Kruskal-Wallis, $H = 11.43$, $p = 0.003$, df = 2). The mean proportion of corpora lutea in the High concentration was less than those in the Control treatment and the Low concentration.
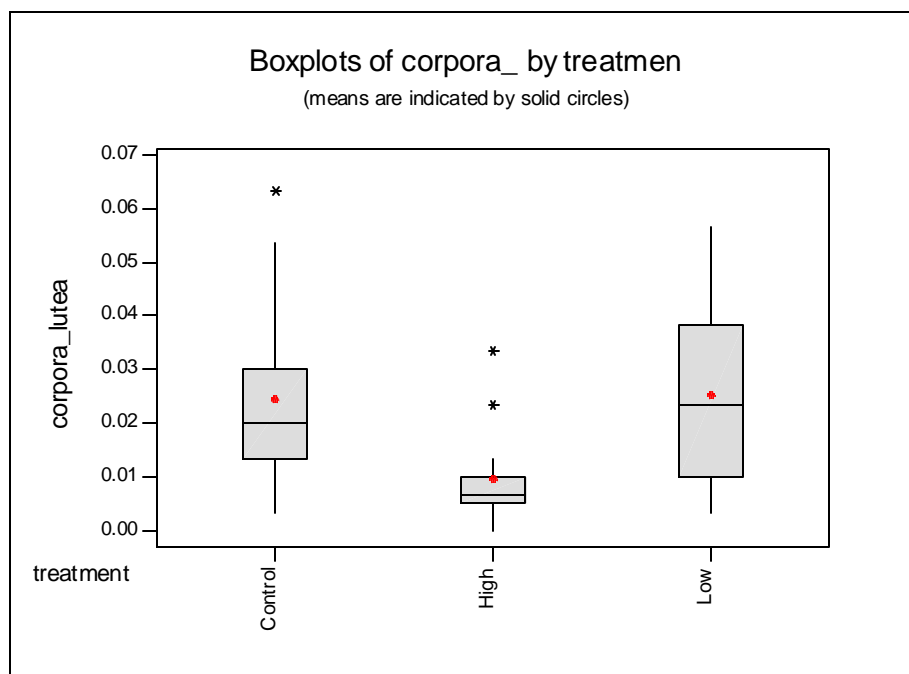
Figure E21T FHist-3. Boxplot of proportion of corpora lutea per 300 follicles by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and asterisks represent probable outliers.

Male Gonad Histology

Testes Staging by Microscopic Field —Testes from males exposed to trenbolone during the EPA 21-day trenbolone assay were examined to determine the general developmental condition. Males in all treatments had well-developed testes with most showing Stage 4 and Stage 5 development (see Methods for description of developmental stages). All of the 96 microscopic fields examined in the 8 Control treatment males showed Stage 4 (79 fields) or Stage 5 (17 fields) development. All of the 120 microscopic fields examined in the 10 Low-concentration treatment males showed Stage 4 (99 fields) or Stage 5 (21 fields) development. All of the 72 microscopic fields examined in the 6 High-concentration treatment males showed Stage 4 (48 fields) or Stage 5 (24 fields) development. Statistical analysis of the mean staging from 12 microscopic fields per fish revealed no significant differences in testicular development among treatments (Kruskal-Wallis, $H = 3.59$, $p = 0.166$, df = 2).

Quantitative Testicular Staging—One hundred cells in each of three sections per male were examined to quantitatively determine the developmental condition of the testes. The developmental stage of the Control treatment testes ranged from Stage 1 to Stage 5, whereas testes in males from the Low and High-concentration treatments showed Stage 2a to Stage 5 development (**Figure E21T MHist-1**). Variability within treatments for each stage was very high as indicated by CVs that ranged as high as 283% (**Table E21T MHist-1**). Although statistical analyses showed that there was a significant difference among treatments in the proportion of cells in developmental Stage 2a and Stage 3b, there were no significant differences among treatments in the proportion of cells in developmental Stages 1, 2b, 3a, 4, and 5 (**Table E21T MHist-1**). Therefore, there was no consistent pattern of significant difference associated with trenbolone dose.

Table E21T MHist-1. Descriptive statistics of the proportion of testes cells in each developmental stage for males from the EPA 21-day trenbolone assay and results of the Kruskal-Wallis Test (df = 2) comparing treatments.

|	Control (n = 8)			Low (n = 10)			High (n = 6)			Kruskal-Wallis	
Stage	Mean	Stdev	CV	Mean	Stdev	CV	Mean	Stdev	CV	*H*	*p*
1	0.0004	0.001	283%	0	0	–	0	0	–	2.00	0.368
2a	0.007	0.003	38%	0.004	0.005	125%	0.002	0.004	182%	6.05	0.049*
2b	0.016	0.009	54%	0.016	0.014	88%	0.014	0.016	114%	0.19	0.910
3a	0.117	0.068	58%	0.156	0.107	69%	0.102	0.077	76%	1.19	0.550
3b	0.280	0.063	22%	0.202	0.112	55%	0.114	0.065	57%	9.44	0.009**
4	0.213	0.081	38%	0.147	0.097	66%	0.115	0.053	46%	4.38	0.112
5	0.367	0.170	46%	0.475	0.252	53%	0.652	0.191	29%	4.84	0.089
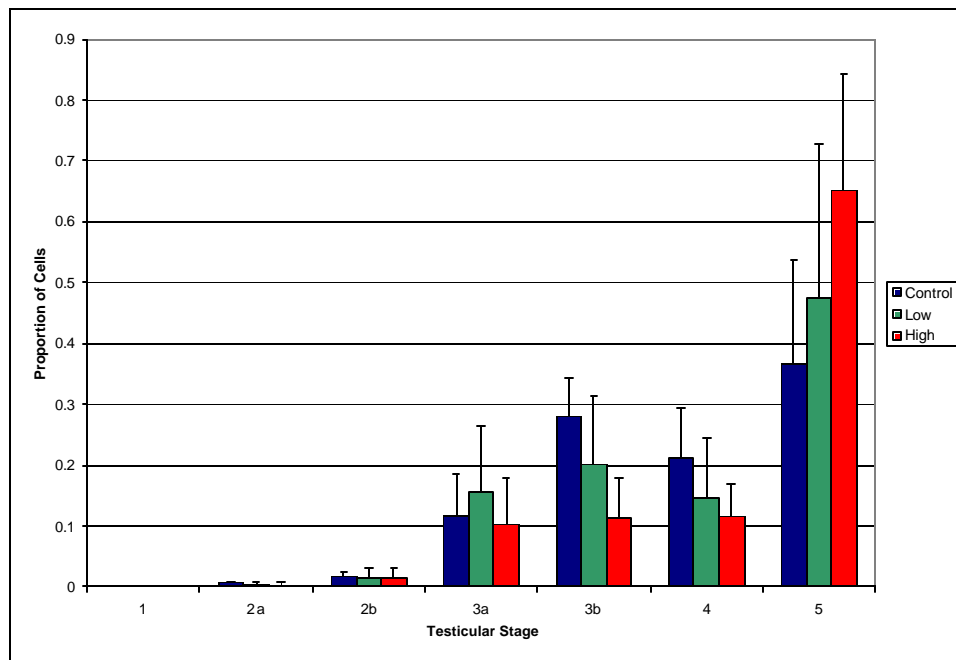
Figure E21T MHist-1. Frequency histogram showing the quantitative developmental staging of testes for each treatment of the EPA 21-day trenbolone assay. For each treatment, the columns represent the grand mean proportion of cells in each stage and the bars represent the standard deviation.

Tubule Diameter—The average diameter of the seminiferous tubules of males from the Control treatment ranged from 78.6 µm to 160.3 µm (**Figure E21T MHist-2**). Tubule diameters of males from the two test concentrations ranged from 88.1 µm to 215.8 µm. No significant differences in the mean tubule diameter per treatment (**Table E21T MHist-2**) were detected (Kruskal-Wallis, $H = 3.33$, $p = 0.190$, df = 2). The achieved power for this endpoint was 33%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 17 (**Table E21T MHist-2**).

Table E21T MHist-2. Summary statistics, detectable difference, and power estimates for male seminiferous tubule diameter data for the EPA 21-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	114.4	28.0	24%	33%	17
low	10	121.4	28.1	23%		
high	6	147.9	37.0	25%		

[1] Calculated from natural log transformed data; with sample size = 2.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
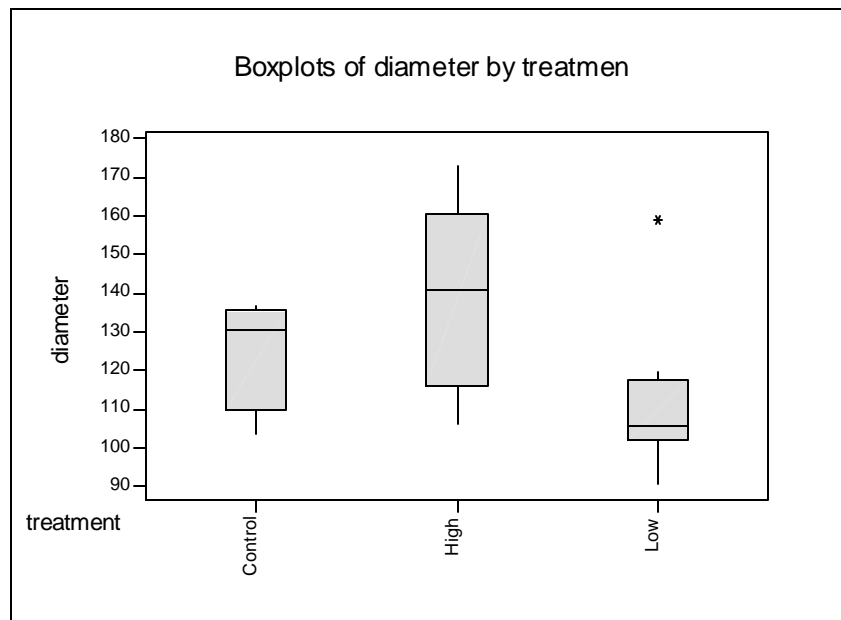

Figure E21T MHist-2. Boxplot of male seminiferous tubule diameter (mm) by treatment for the EPA 21-day trenbolone assay. The box re presents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

> **Observations** —One male in the Control treatment showed early stage cells in the lumen of some tubules. No interstitial Sertoli cell proliferation and no Leydig cell proliferation was observed for any treatment. No testicular atrophy was recorded and no ovatestes were observed for any treatment.

Vitellogenin
> Vitellogenin concentrations in most Control treatment females used during the EPA 21-Day trenbolone assay ranged from 329,100 ng/mL to 1,434,000 ng/mL (**Figure E21T VTG-1**). Two control-treatment females had vitellogenin concentrations of 3,443,500 ng/mL and

4,162,000 ng/mL. Among females exposed to the two trenbolone concentrations, vitellogenin concentrations ranged from 0 ng/mL (not detected) to 3,184,500 ng/mL. Significant differences in the mean vitellogenin concentration per treatment (**Table E21T VTG-1**) were detected (Kruskal-Wallis, $H = 21.37$, $p = <0.001$, df = 2). Vitellogenin concentrations in Control-treatment females and low-concentration females were significantly greater than those in females exposed to the High trenbolone concentrations. The achieved power for this endpoint was 100%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 5 (**Table E21T VTG-1**).

Table E21T VTG-1. Summary statistics, detectable difference, and power estimates for female vitellogenin concentrations (ng/mL) for the EPA 21-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	13	1,289,831	1,164,016	90%	99%	7
low	14	1,221,589	1,093,574	90%		
high	15	113,444	270,367	238%		

[1] Calculated from natural log transformed data; with sample size = 13.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
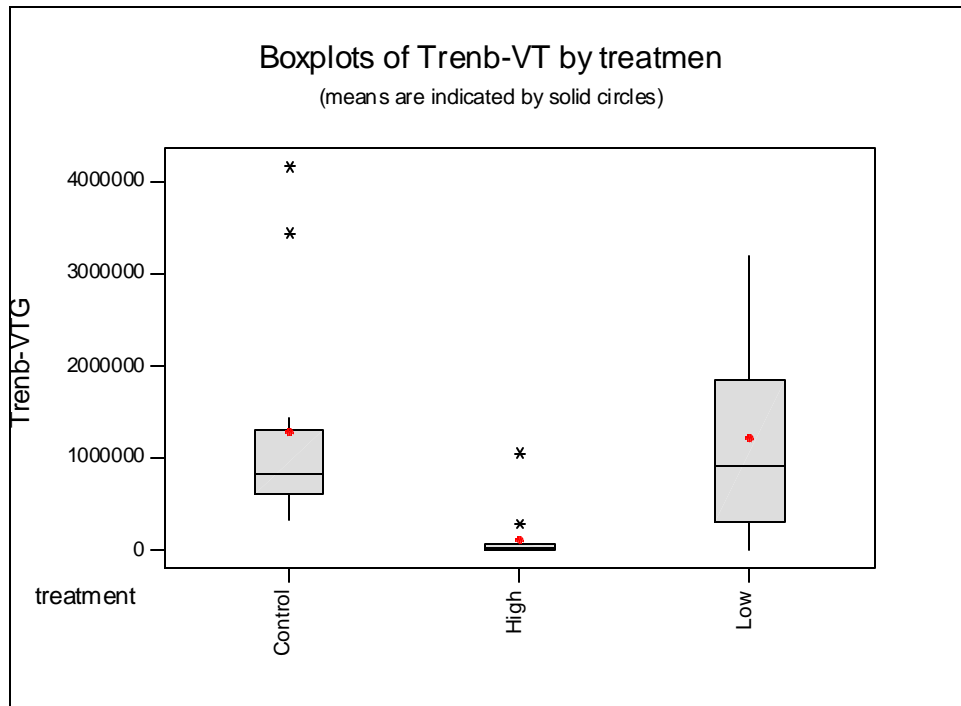

Figure E21T VTG-1. Boxplot of female vitellogenin concentration (ng/mL) by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and asterisks represent probable outliers.

Vitellogenin concentrations in Control treatment males used during the EPA 21-Day trenbolone assay ranged from 0 ng/mL (not detected) to 8,368 ng/mL (**Figure E21T VTG-1**). Among males exposed to the two trenbolone concentrations, vitellogenin concentrations ranged from 0 ng/mL (not detected) to 185,700 ng/mL. No significant differences in the mean vitellogenin concentration per treatment (**Table E21T VTG-1**) were detected (Kruskal-Wallis, $H = 0.65$, $p = 0.722$, df = 2). The achieved power for this endpoint was 8%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 140 (**Table E21T VTG-1**).

Table E21T VTG-1. Summary statistics, detectable difference, and power estimates for male vitellogenin concentrations (ng/mL) for the EPA 21-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	2,919	3,607	124%	8%	140
low	8	43,439	76,533	176%		
high	7	6,983	11,964	171%		

[1] Calculated from natural log transformed data; with sample size = 7.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
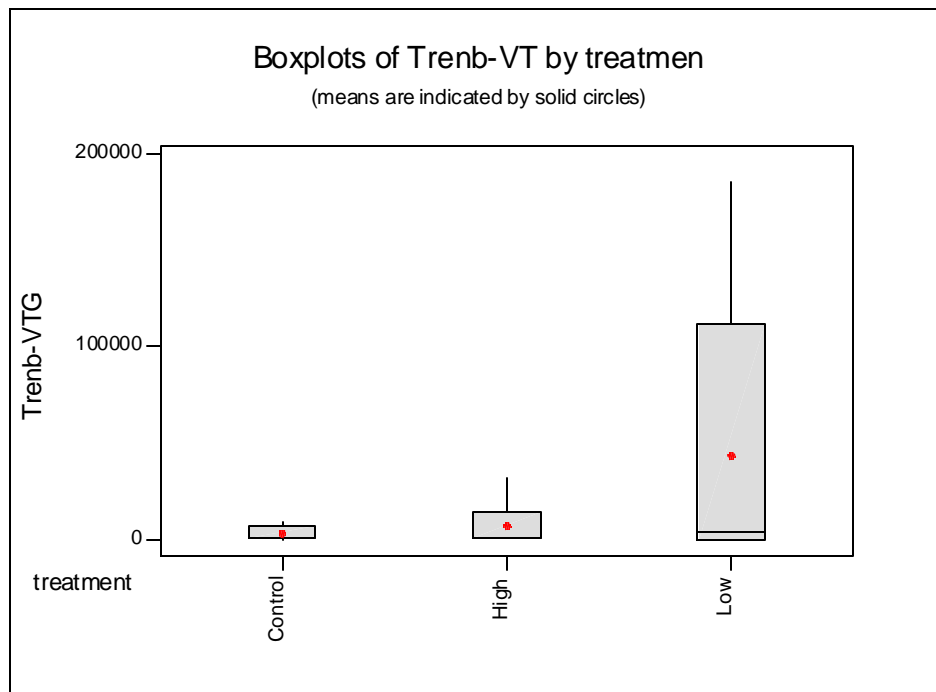

Figure E21T VTG-1. Boxplot of male vitellogenin concentration (ng/mL) by treatment for the EPA 21-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Plasma Steroid Concentrations

3.2.3 EPA 14-day Assay for trenbolone

The EPA 14-day trenbolone assay was conducted from February 4, 2003 to February 11, 2003 (prevalidation test) and from February 11, 2003 to February 25, 2003 (validation test).

Chemistry

Trenbolone concentrations were measured in each replicate on Day 6 and 14 for the High concentration and Day 2, 6, and 14 for the Low concentration (**Figure E14M Chem-1**). The mean percent difference of the High and Low concentrations achieved during the test versus the target concentrations were 22% and 45%, respectively (**Table E14M Chem-1**).

Table E14T Chem-1. Summary statistics for trenbolone concentrations (mg/L) during the EPA 14-Day assay.

		Treatment	
	Control	**Low**	**High**
Target	0.051 [1]	0.1	1
mean	0.051 U	0.055 [2]	0.78
stdev	0	0.022	0.08
min	0.051 U	0.021	0.67
max	0.051 U	0.093	0.89
median	0.051 U	0.055	0.78
Q1	0.051 U	0.046	0.72
Q3	0.051 U	0.066	0.83
Mean % Δ	–	45%	22%

[1] Equal to the method detection limit (MDL).
[2] Concentrations in some replicates were estimated at levels less than the MDL.
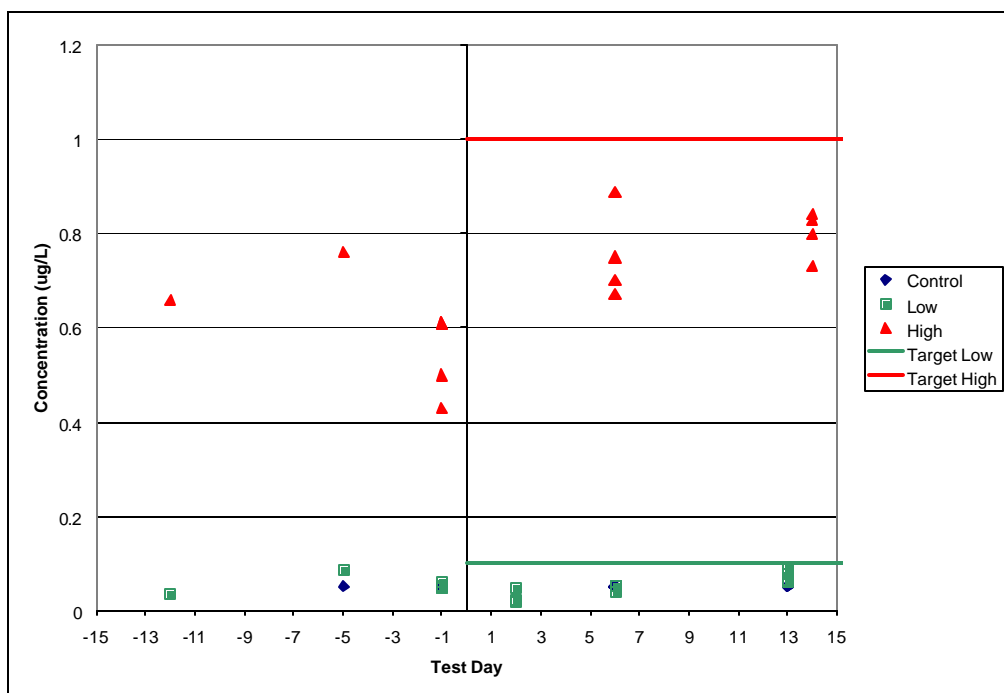U Not detected at concentrations greater than the MDL.

Figure E14T Chem-1. trenbolone concentrations measured in the High and Low-concentration treatments versus the target concentrations (1.0 ㎍/L and 0.1 ㎍/L) during the EPA 14-day trenbolone assay. Control values are plotted at the MDL.

Test Conditions Summary

      All water quality, water chemistry, and light intensity parameters were within the acceptable ranges during the EPA 14-day trenbolone prevalidation and validation tests (Table E14T WQ-1). Flow rates were monitored regularly, and were adjusted if they were observed to be out of the target range (see Methods). Hardness, alkalinity, and ammonia were not measured during the pre-exposure assay in tanks that would be used for the Low concentration during the exposure assay.

Table E14T WQ-1. Summary of water quality, water chemistry, flow rate, and light intensity data for the EPA 14-day trenbolone assay.

	Temperature (°C)		Dissolved Oxygen (mg/L)		pH		Hardness (mg/L)		Alkalinity (mg/L)		Ammonia (mg/L)		Light Intensity (lux)			
	min	max	min	max	min	max	min	max	min	max	min	max	min	max	mean	SD
Pre-exposure Assay																
Target Range	24	26	>4.9		6.5	9.0	>140		>20		<0.5		540	1080		
Control	24.0	24.9	6.8	8.2	7.0	8.3	170	170	150	150	0.1	0.1	679	834	780	58
Low	24.0	25.2	6.4	8.2	7.4	8.3	–	–	–	–	–	–	742	918	814	59
High	24.0	25.0	6.7	8.1	7.7	8.3	170	170	150	150	0.1	0.1	730	863	814	59
Exposure Assay																
Target Range	24	26	>4.9		6.5	9.0	>140		>20		<0.5		585	1060	906	154
Control	24.1	25.2	6.4	7.6	7.4	8.2	170	221	135	160	0.1	0.5	707	931	828	98
Low	24.1	25.2	6.4	7.8	7.5	8.2	170	204	140	160	0.1	0.3	551	932	806	175
High	24.0	25.1	6.9	7.9	7.5	8.2	170	204	130	150	0.1	0.2	763	944	836	77

Survival

All males in all treatments and all females in the Control and High-concentration treatments survived the EPA 14-day trenbolone assay. Two females in the Low-concentration treatment died during the assay (88% survival).

Fecundity

Total Fecundity—Variability among treatments in the total number of eggs produced during the EPA 14-day trenbolone assay was very high (**Figure E14T Fec-1**). Total counts in the Control treatment ranged from 1,951 eggs to 3,227 eggs. Total counts for three replicates of the High trenbolone concentration treatment were similar, ranging from 280 eggs to 346 eggs, whereas the total number of eggs produced in the fourth replicate was 666. It is also notable that the primary egg production in three High trenbolone concentration replicates occurred only on Day 1 (one replicate yielded 30 eggs on Day 12). Egg production in the fourth High trenbolone concentration occurred only on Days 2 and 3. The minimum and maximum average production among replicates in the Low trenbolone concentration ranged from 2,339 eggs to 3,140 eggs. Statistical analysis of square-root transformed egg counts showed significant among-treatment differences in the mean cumulative number of eggs produced per treatment (**Table E14T Fec-1**) (1-way ANOVA, $F = 63.99$, $p < 0.001$, df = 2, 9). Dunnett's comparison identified significant differences in mean egg production between the High trenbolone concentration and the Control treatment, and also between the High and Low concentrations. The Low concentration and the Control treatment fecundity values were not statistically different. The achieved power for this assay was 100%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 2 (**Table E14M Fec-1**).

Table E14T Fec-1. Summary statistics, detectable difference, and power estimates for fecundity data for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	4	2654	600	23%	100%	2
low	4	2689	348	13%		
high	4	400	180	45%		

[1] Calculated from square-root transformed data; with sample size = 4.

[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on square-root transformed data.
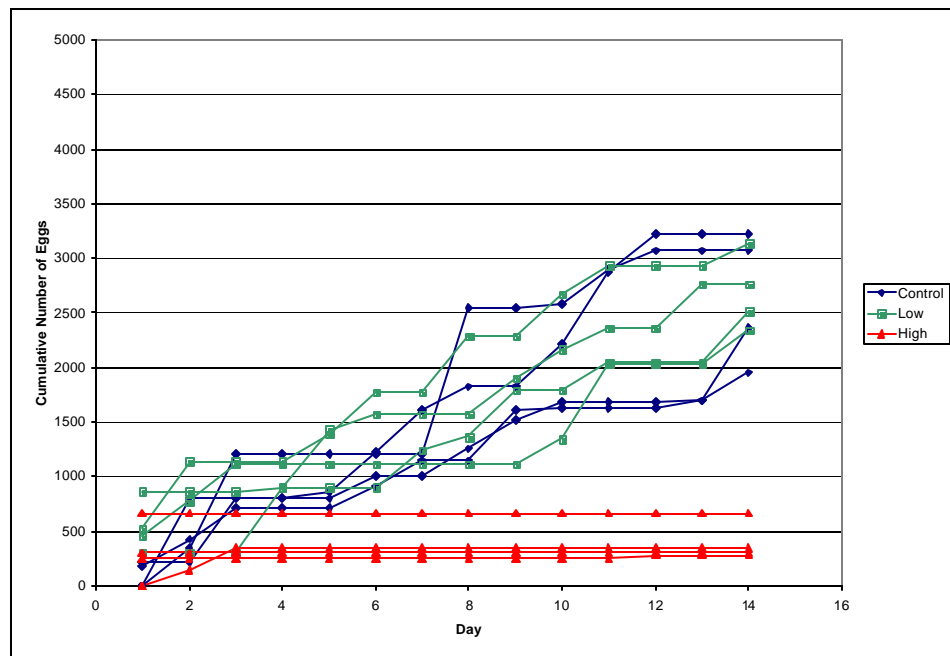

Figure E14T Fec-1. Total egg production by replicate per treatment for the EPA 14-Day trenbolone assay.

Fecundity per Female Reproductive Day—During the EPA 14-day trenbolone assay, the maximum number of female reproductive days was achieved for the Control treatment and the High concentration, whereas 55 female reproductive days were achieved in the Low concentration (**Table E14T Fec-2**). The number of eggs produced per female reproductive day in the Control treatment varied from 34.8 eggs to 57.6 eggs and from 41.8 eggs to 58.1 eggs in the Low concentration (**Figure E14T FEC-2**). For the High concentration, the number of eggs produced per female reproductive day ranged from 5.0 eggs to 11.9 eggs, with fish in three of the replicates producing fewer than 7 eggs per day (5.0, 5.5, 6.2). A significant difference in the mean number of eggs produced per female reproductive day among treatments was detected (Kruskal-Wallis, $H = 7.42$, $p = 0.024$, df = 2). Fewer eggs were produced per female reproductive day in the High concentration than in either the Control treatment or the Low concentration. There was no difference in the mean number of eggs laid per female reproductive day between the Control treatment and the Low concentration. The achieved power for this assay

was 100%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 2 (**Table E14M Fec-2**).

Table E14T Fec-2. Summary statistics, detectable difference, and power estimates for fecundity per female reproductive day for the EPA 14-day trenbolone assay.

Level	Mean Number of Reproductive Days [1]	N	Mean	Stdev	CV	Achieved Power [2]	Sample Size Required [3]
control	56.0	4	47.4	10.7	23%	100%	2
low	55.0	4	48.5	7.1	15%		
high	56.0	4	7.2	3.2	45%		

[1] Maximum number = 56.
[2] Calculated from natural log transformed data; with sample size = 4.
[3] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
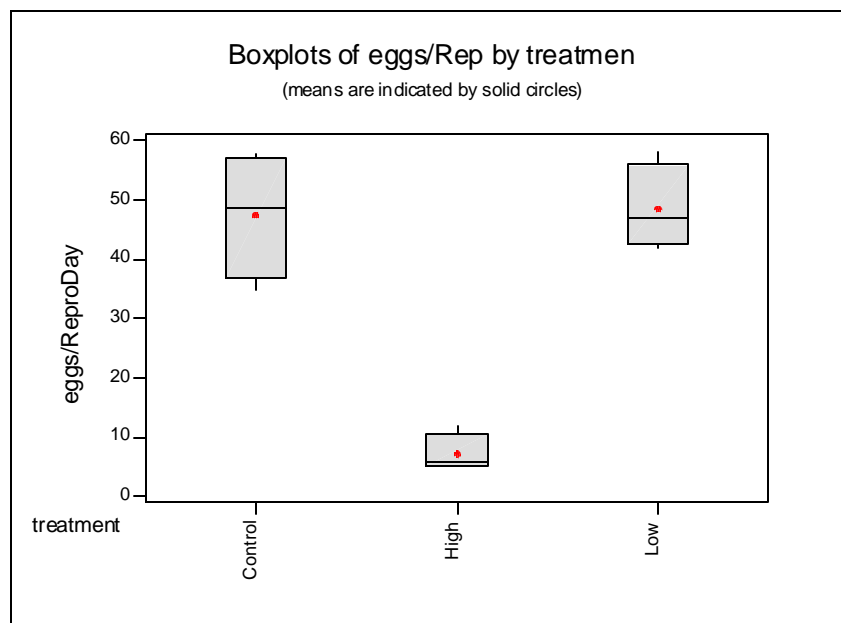

Figure E14T FEC-2. Boxplot of the number of eggs produced per female reproductive day by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Eggs on Tiles/Dishes—The mean number of eggs laid on the tiles among the treatments during the EPA 14-day trenbolone assay varied from 350 eggs for the High concentration to 2,432 eggs for the Control treatment (**Table E14T FEC-3**). The number of eggs on dishes

ranged from 50 eggs for the High concentration to 222 eggs for the Control treatment. Because of the variability in the total number of eggs laid per treatment, the proportional difference in the number of eggs on dishes versus those on tiles [1−(# eggs on dishes ÷ # eggs on tiles)] was calculated. The proportional difference ranged from 0.78 (one High concentration replicate) to 0.96 (one low concentration replicate) (**Figure E14T FEC-3**). There were no significant differences in this mean proportional difference among treatments (Kruskal-Wallis, $H = 3.23$, $p = 0.199$, df = 2).

Table E14T FEC-3. Summary statistics for the numbers of eggs laid on tiles versus those in dishes for the EPA 14-day trenbolone assay.

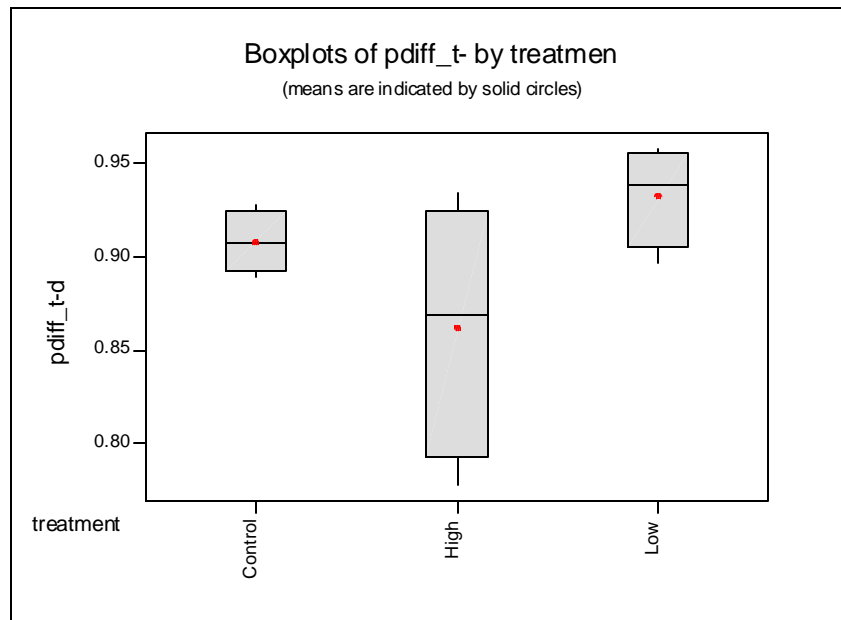|		Tile			Dish			Mean Proportional Difference
Level	**N**	**Mean**	**Stdev**	**CV**	**Mean**	**Stdev**	**CV**	
control	4	2432	558	23%	222	58	26%	0.91
low	4	2526	376	15%	162	44	27%	0.93
high	4	350	150	43%	50	34	68%	0.86

Figure E14T FEC-3. Boxplot of the proportional difference in the number of eggs laid on tiles versus those in dishes by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Fertilization Success

Total Fertilization—The total (tiles + dishes) fertilization-success rates for most treatment replicates during the EPA 14-day trenbolone assay were high, ranging from 0.981 (one High concentration replicate) to 1.00 (one Control treatment replicate) (**Figure E14T FER-1**). A significant difference in mean fertilization-success rate s (**Table E14T FER-**1) among treatments was detected (Kruskal-Wallis, $H = 7.65$, $p = 0.022$, df = 2). The fertilization-success rate was lower in the High concentration than in the Control treatment or the Low concentration. There

was no difference in the fertilization-success rate between the Control treatment and the Low concentration. The achieved power for this assay was 89%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 4 (**Table E14T FER-1**).

Table E14T FER-1. Summary statistics, detectable difference, and power estimates for the proportion of eggs fertilized for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	4	0.999	0.001	0%	89%	4
low	4	0.998	0.002	0%		
high	4	0.990	0.007	1%		

[1] Calculated from arcsine square-root transformed data; with sample size = 4.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
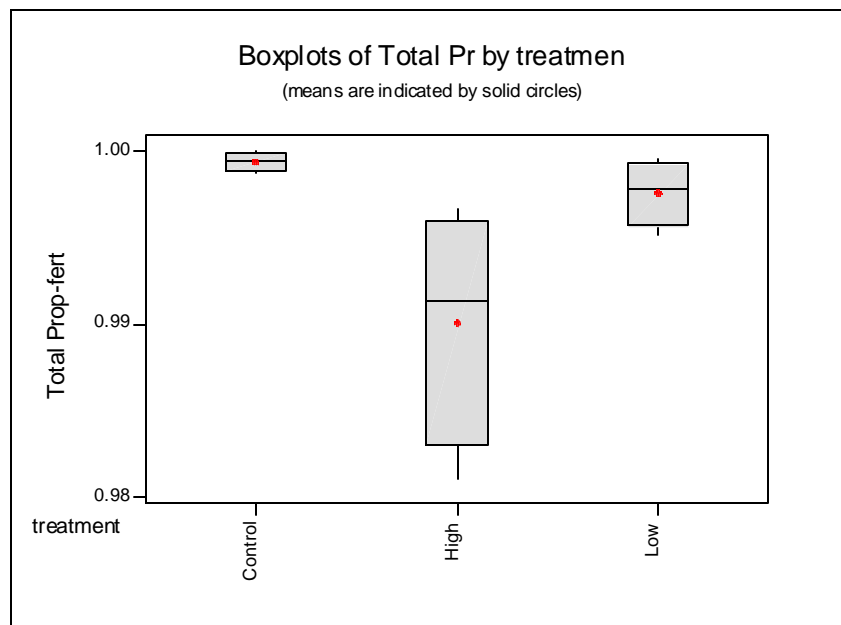

Figure E14T FER-1. Boxplot of the proportion of eggs fertilized by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Fertilization of Eggs on Tiles and Dishes—The fertilization-success rates for all treatment replicates for eggs laid on tiles during the EPA 14-day trenbolone assay were high, ranging from 0.978 (one High concentration replicate) to 1.00 (two Control treatment replicates, one High and one Low concentration replicate) (**Figure E14T FER-2**). No significant differences in mean fertilization-success rates for eggs laid on tiles (**Table E14T FER-2**) among treatments were detected (Kruskal-Wallis, $H = 3.16$, $p = 0.206$, df = 2). The fertilization-success rates for all treatment replicates for eggs laid on dishes during the assay were high, ranging from 0.933 (one High concentration replicate) to 1.00 (several replicates; all treatments) (**Figure E14T**

FER-3). No significant differences in mean fertilization-success rates (**Table E14T FER-3**) among treatments were detected (Kruskal-Wallis, $H = 0.04$, $p = 0.981$, df = 2).

Table E14T FER-2. Summary statistics for the proportion of eggs on tiles versus those in dishes that were fertilized for the EPA 14-day trenbolone assay.

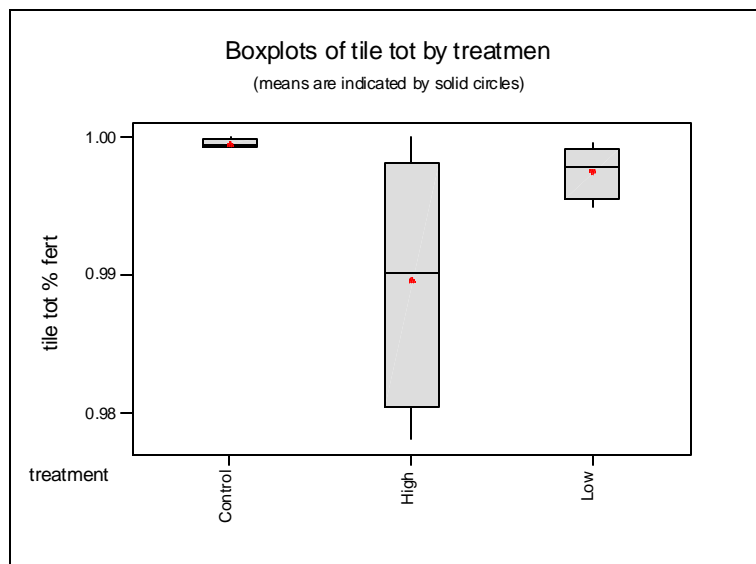|		Tiles			Dishes		
Level	**N**	**Mean**	**Stdev**	**CV**	**Mean**	**Stdev**	**CV**
control	4	1.00	0.00	0%	1.00	0.00	0%
low	4	0.998	0.002	0%	0.999	0.002	0%
high	4	0.990	0.009	1%	0.983	0.033	3%

Figure E14T FER-2. Boxplot of the proportion of eggs on tiles that were fertilized by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.
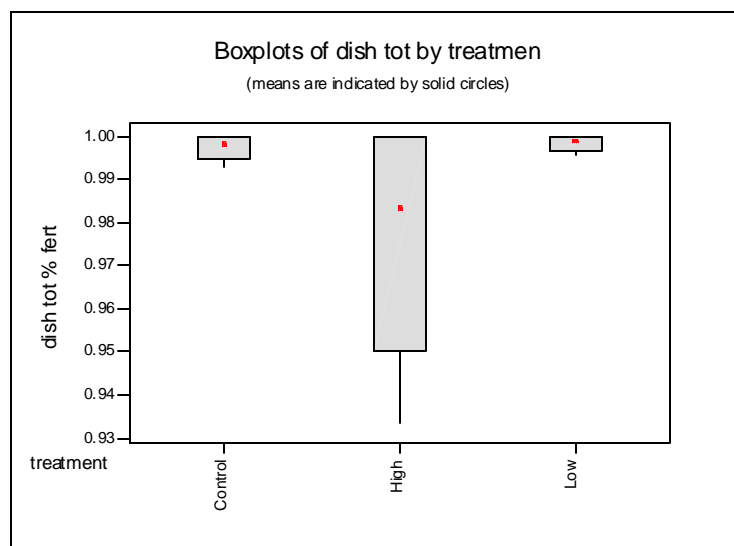
Figure E14T FER-3. Boxplot of the proportion of eggs in dishes that were fertilized by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Hatchability and Larval Development

Eggs were collected during the 14-day pre-exposure period for the evaluation of hatchability. The mean proportion of fertilized eggs that hatched was 0.98 (sd = 0.03) for Control treatment, 0.94 (sd = 0.03) for the Low concentration, and 0.97 (sd = 0.03) for the High concentration. Among the tanks evaluated during the pre-exposure period, but that were not used in the 21-Day assay, the mean proportion of fertilized eggs that hatched was 0.75 (sd = 0.23), which was significantly lower than those for the other treatments (Kruskal-Wallis, $H = 9.59$, $p = 0.022$, df = 3).

Eggs were collected during the EPA 14-day trenbolone assay for hatchability analysis. Fish in the High concentration did not produce eggs during that period, therefore, the High concentration was not included in the analysis. The proportion of fertilized eggs that hatched ranged from 0.96 to 1.00 in the Control treatment and was 1.00 for all replicates of the Low concentration (**Figure E14T HAT-1**). There were no significant differences between the two treatments in the proportion of eggs that hatched (Kruskal-Wallis, $H = 2.29$, $p = 0.131$, df = 2). The achieved power for this endpoint was 11%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 13 (**Table E14T HAT-1**).

Table E14T HAT-1. Summary statistics, detectable difference, and power estimates for the proportion of fertile eggs that hatched for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	4	0.98	0.02	2%	11%	13
low	4	1.00	0	0%		
high	0	–	–	–		

[1] Calculated from arcsine square-root transformed data; with sample size = 4.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
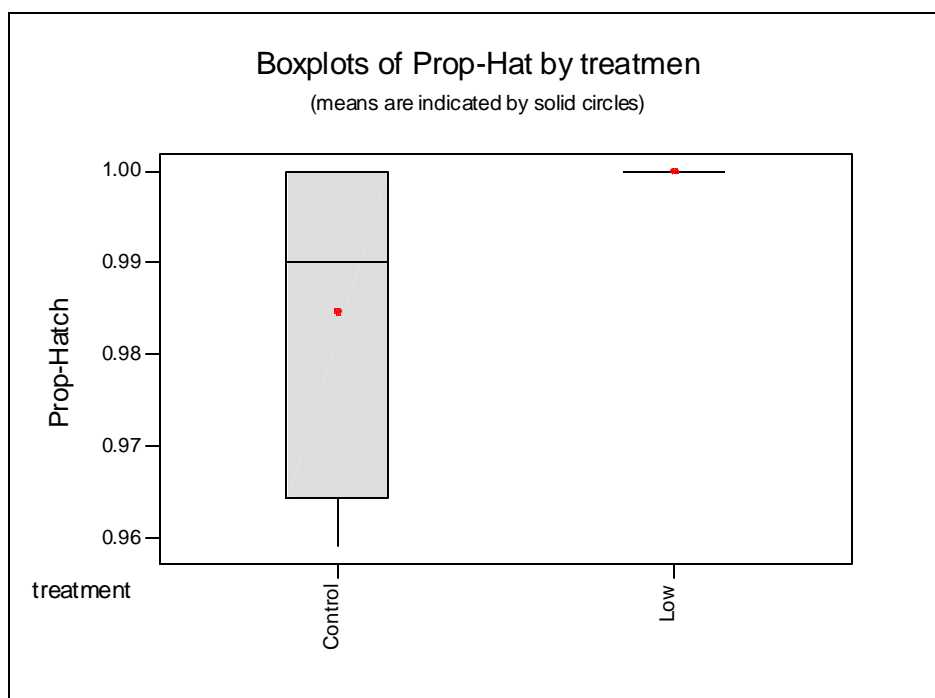
Figure E14T HAT-1. Boxplot of the proportion of fertile eggs that hatched by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Eggs were collected during the pre-exposure period for the evaluation of larval development. The mean proportion of larvae that developed normally (i.e., that showed no morphological abnormalities) was 0.98 (sd = 0.02) for Control treatment. The mean proportion of normal larvae in the remaining treatments was 0.98 (sd = 0.02) in the Low concentration and 0.99 (sd = 0.01) in the High concentration. Among the tanks evaluated during the pre-exposure period, but that were not used in the 21-Day assay, the mean proportion of normal larvae was 0.98 (sd = 0.03). There were no significant differences among treatments in the proportion of eggs that hatched (Kruskal-Wallis, $H = 0.84$, $p = 0.840$, df = 3).

Eggs were collected on Days 7 through 10 during the EPA 14-day trenbolone assay for the evaluation of larval development. Fish in the High concentration did not produce eggs during that period, therefore, the High concentration was not included in the analysis. The proportion of larvae that developed normally (i.e., that showed no morphological abnormalities) ranged from 0.96 to 1.00 in the Control treatment and from 0.78 to 1.00 for the Low concentration (**Figure E14T HAT-2**). There were no significant differences among treatments in the proportion of eggs that hatched (Kruskal-Wallis, $H = 0.79$, $p = 0.375$, df = 2). The achieved power for this endpoint was 8%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 29 (**Table E14T HAT-2**).

Table E14T HAT-2. Summary statistics, detectable difference, and power estimates for the proportion of normal larvae for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	4	0.98	0.02	2%	8%	29
low	4	0.92	0.10	11%		
high	0	–	–	–		

[1] Calculated from arcsine square-root transformed data; with sample size = 4.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
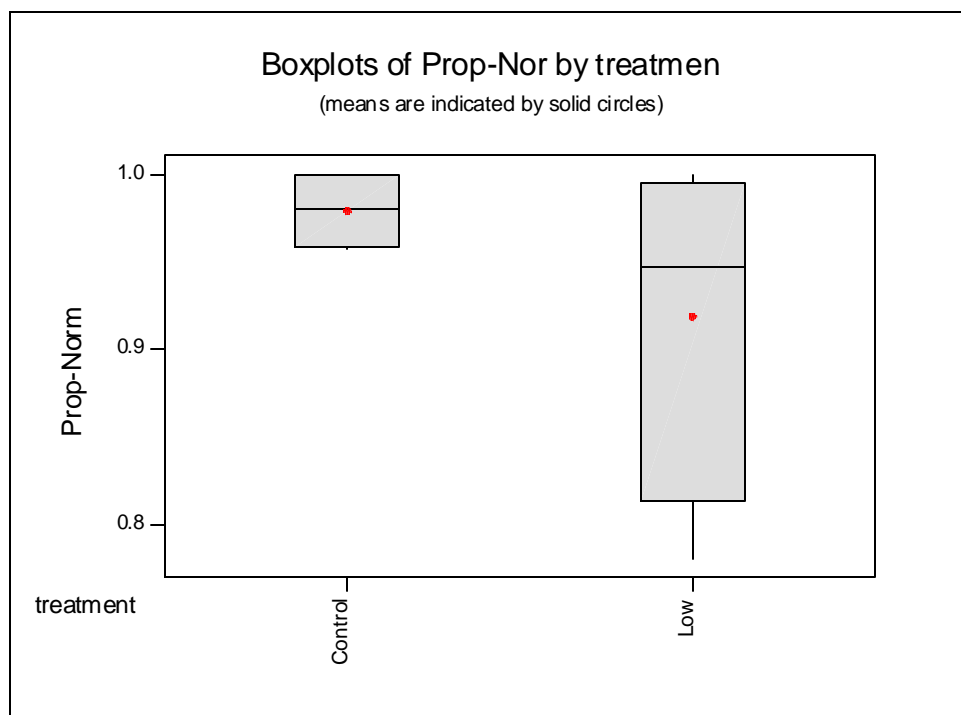

Figure E14T HAT-2. Boxplot of the proportion of normal larvae by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Body Weight

The body weight of females used in the EPA 14-day trenbolone assay ranged from 1.0 g to 2.9 g. A significant difference in the mean female body weight per treatment was detected (Kruskal-Wallis, $H = 10.77$, $p = 0.005$, df = 2). The test indicated that the mean female body weight for females in the Low-concentration treatment was less than those for the other two treatments. The body weight of males used in the EPA 14-day trenbolone assay ranged from 2.4 g to 4.7 g. There were no significant differences in mean body weight among treatments (Kruskal-Wallis, $H = 3.72$, $p = 0.156$, df = 2).

Gonadosomatic Index

The range of GSI values calculated for females in the all treatments varied from two- to three-fold (**Figure E14T GSI-1**), and the overall variability within the treatment was moderate (CVs = 23% –27%; **Table E14T GSI-1**). GIS values for fish in the Control treatment ranged from 6.4 to 14.3. The highest female GSI value was about 24 (one fish in the High concentration). No other fish had a GSI value greater than 18.6. A significant difference in the mean GSI value per treatment (**Table E14T GSI-1**) was detected (Kruskal-Wallis, $H = 9.84$, $p = 0.007$, df = 2). The test indicated that the mean GSI for females in the Control treatment was less than those for the other two treatments. However, the mean GSI values calculated for the Low and High trenbolone concentrations were not statistically different. Because female body weights were not different between the two treatments (previous section), the gonads of females in the Low concentration were proportionally larger (i.e., hypertrophied) compared to those in the High concentration. The achieved power for this endpoint was 78%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 15 (**Table E14T GSI-1**).

Table E14T GSI-1. Summary statistics, detectable difference, and power estimates for female gonadosomatic index data for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	16	10.5	2.6	25%	78%	15
low	14	13.7	3.1	23%		
high	16	14.3	3.8	27%		

[1] Calculated from arcsine square-root transformed data; with sample size = 14.

[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
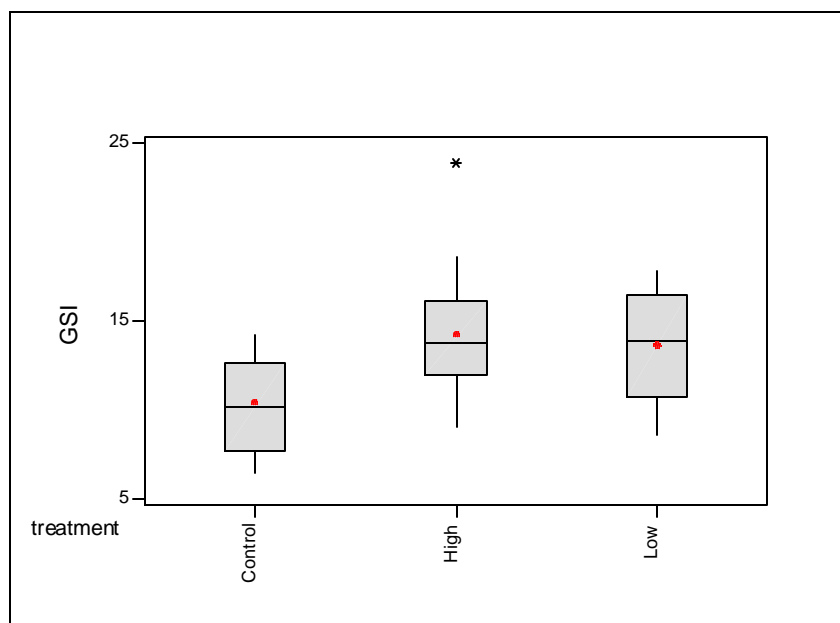
Figure E14T GSI-1. Boxplot of female GSI by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

The range of most GSI values calculated for males during the EPA 14-day trenbolone assay, was small, ranging from 0.7 to 1.8 (**Figure E14T GSI-2**), which approximates the typical range for reproductively-active male fathead minnows. The highest and lowest male GSI values were 2.2 (one fish the High concentration) and 0.7 (two fish in the Control treatment), respectively. A significant difference in the mean GSI value per treatment (**Table E14T GSI-2**) was detected (Kruskal-Wallis, $H = 7.66$, $p = 0.022$, df = 2). The test indicated that the mean GSI for males in the High concentration was greater than those for the other two treatments. The achieved power for this endpoint was 71%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 10 (**Table E14T GSI-2**).

Table E14T GSI-2. Summary statistics, detectable difference, and power estimates for male gonadosomatic index data for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	1.04	0.28	27%	71%	10
low	8	1.02	0.33	33%		
high	8	1.56	0.39	25%		

[1] Calculated from arcsine square-root transformed data; with sample size = 8.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
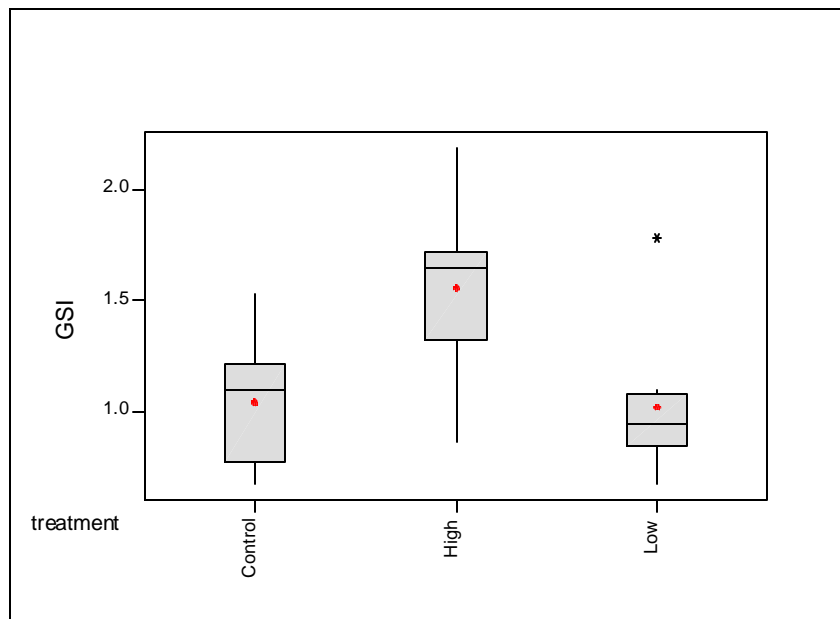

Figure E14T GSI-2. Boxplot of male GSI by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

Female Gonad Histology

General Ovary Staging—Statistical analysis of the mean ovarian staging from 12 microscopic fields per female in the EPA 14-day trenbolone assay revealed no significant difference among treatments (Kruskal-Wallis, $H = 2.21$, $p = 0.331$, df = 2).

Quantitative Ovarian Staging—One hundred cells in each of three sections per female were examined to quantitatively determine the developmental stage of the ovaries. Ova from fish in all treatments ranged from Stage 1a to Stage 5 (see Methods for a description of the stages (**Figure E14T FHist-1**). Variability within treatments for each stage was very high as indicated by CVs that ranged as high as 400% (**Table E14T FHist-1**). Although statistical analyses

showed that there was a significant difference among treatments in the proportion of cells in developmental Stage 3 and Stage 4, there were no significant differences among treatments in the proportion of cells in the developmental Stages 1a, 1b, 2, and 5 (**Table E14T FHist-1**). The proportion of cells in developmental Stage 3 in the High concentration was significantly greater than that in the Control treatment and the Low concentration. The proportion of cells in developmental Stage 4 in the High concentration was significantly lower than that in the Control treatment and the Low concentration. Thus, it can be concluded that the trenbolone High concentration inhibited the development of ova from the early to late vitellogenic stages (*i.e.* from stage 3 to stage 4).

Table E14T FHist-1. Descriptive statistics of the proportion of ovarian cells in each developmental stage for females from the EPA 14-day trenbolone assay and results of the Kruskal-Wallis Test (df = 2) comparing treatments.

| Stage	Control (n = 16)			Low (n = 15)			High (n = 16)			Kruskal-Wallis	
	Mean	Stdev	CV	Mean	Stdev	CV	Value	Stdev	CV	*H*	*p*
1a	0.104	0.037	36%	0.096	0.033	35%	0.082	0.023	28%	2.31	0.315
1b	0.321	0.064	20%	0.304	0.057	19%	0.294	0.059	20%	1.56	0.459
2	0.194	0.050	26%	0.174	0.044	25%	0.161	0.042	26%	5.77	0.056
3	0.138	0.037	27%	0.158	0.051	32%	0.291	0.056	19%	29.47	<0.001***
4	0.215	0.091	42%	0.241	0.079	33%	0.130	0.068	52%	14.39	0.001***
5	0.003	0.011	339%	0.001	0.003	300%	0.001	0.003	400%	0.42	0.809
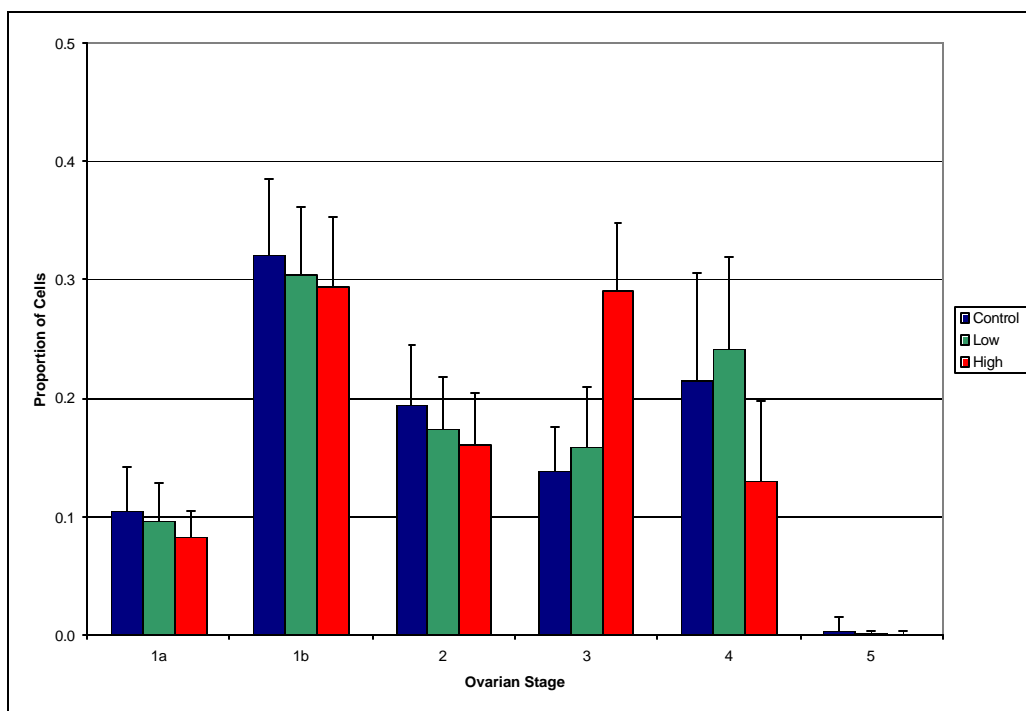
Figure E14T FHist-1. Frequency histogram showing the quantitative developmental staging of ovaries for each treatment of the EPA 14-day trenbolone assay. For each treatment, the columns represent the grand mean proportion of cells in each stage and the bars represent the standard deviation.

      Atretic Follicles—The mean proportion of atretic follicles per 300 follicles (counted per fish) ranged from 0.002 for females in the Control treatment to 0.03 follicles for females in the High concentration (**Figure E14T FHist-2**). There was a significant difference in the proportions of atretic follicles among treatments (Kruskal-Wallis, $H = 23.99$, $p = <0.001$, df = 2). The proportion of atretic follicles in females from the High concentration was significantly greater than that in Control treatment and Low concentration females. This appears to be related to the inhibition of development of ova from stage 3 to stage 4 in the High-concentration treatment.
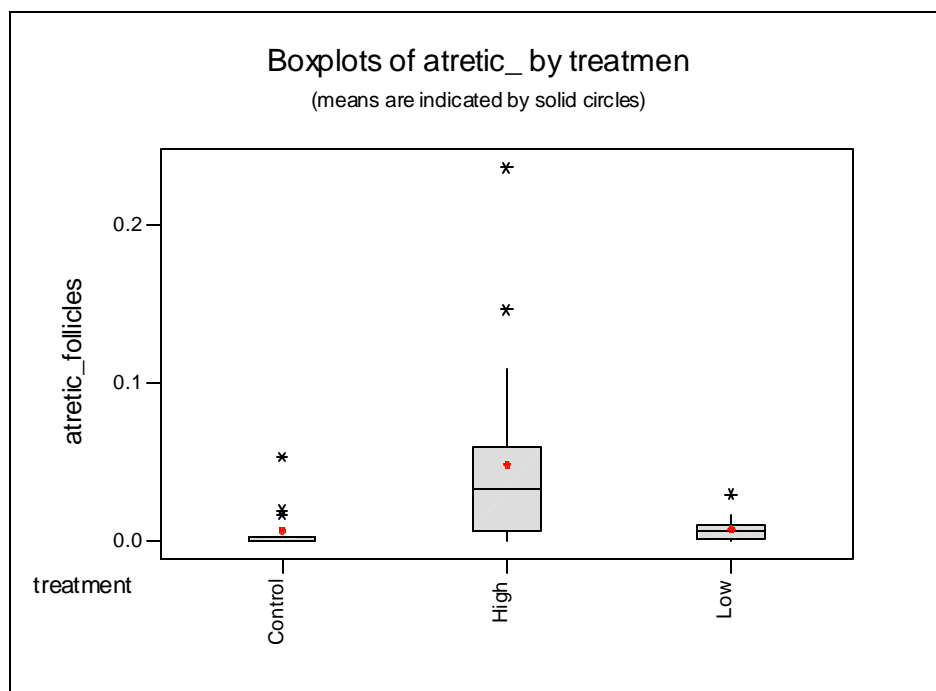
Figure E14T FHist-2. Boxplot of the proportion of atretic follicles per 300 follicles by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and asterisks represent probable outliers.

       Corpora Lutea—The mean proportion of corpora lutea per 300 follicles (counted per fish) ranged from 0.01 for females in the High concentration to 0.02 for females in the Control treatment and the Low concentration (**Figure E14T FHist-3**). There was a significant difference in the proportions of corpora lutea among treatments (Kruskal-Wallis, $H = 7.04$, $p = 0.030$, df = 2). The value for the High concentration was significantly lower than that of the Low-concentration treatment. This may be related to the inhibition of ovarian development in the High concentration, so that over time, while ovarian development tends to be stalled between stages 3 and 4, the corpora lutea in the High dose females are resorbed.

       Observations—One female in the Control treatment had abnormal Stage 3 ova that were more typical of Stage 2, but had vitellogenin granules in the cytoplasm.
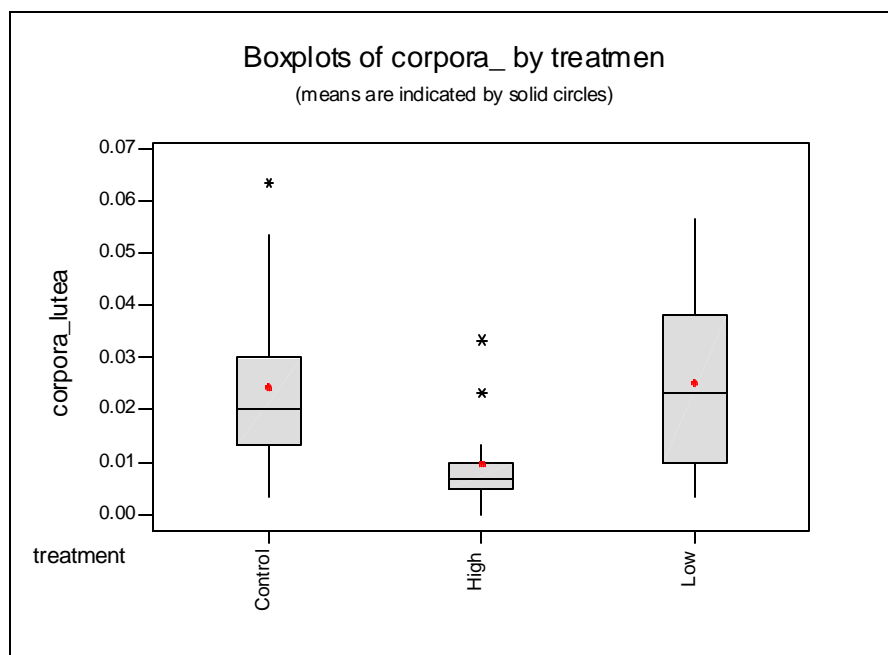
Figure E14T FHist-3. Boxplot of the proportion of corpora lutea per 300 follicles by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and asterisks represent probable outliers.

Male Gonad Histology

      Testes Staging by Microscopic Field—Testes from males exposed to trenbolone during the EPA 14-day trenbolone assay were examined to determine the general developmental condition. Males in all treatments had well-developed testes with most showing Stage 4 and Stage 5 development (see Methods for description of developmental stages). All of the 96 microscopic fields examined in the 8 Control treatment males showed Stage 4 (91 fields) or Stage 5 (5 fields) development. All of the 96 fields examined in the 8 Low-concentration treatment males showed Stage 4 (90 fields) or Stage 5 (6 fields) development. All of the 96 microscopic fields examined in the 8 High-concentration treatment males showed Stage 4 (58 fields) or Stage 5 (38fields) development. Statistical analysis of the mean staging from 12 microscopic fields per fish revealed a significant difference among treatments (Kruskal-Wallis, $H = 12.36$, $p = 0.002$, df = 2). The mean testes stage for males in the High concentration (4.40, sd = 0.29) was significantly greater than that in the Control treatment (4.05, sd = 0.09) and Low-concentration treatment(4.06, sd = 0.10).

      Quantitative Testicular Staging—One hundred cells in each of three sections per male were examined to quantitatively determine the developmental condition of the testes. The developmental stage of the testes from fish in all treatments ranged from Stage 2a to Stage 5 (**Figure E14T MHist-1**). Variability within treatments for each stage was very high as indicated by CVs that ranged as high as 129% (**Table E14T MHist-1**). Although statistical analyses showed that there was a significant difference among treatments in the proportion of cells in developmental Stage 3a, there were no significant differences among treatments in the proportion of cells in developmental Stages 2a, 2b, 3b, 4, and 5 (**Table E14T MHist-1**). Therefore, there was no consistent pattern of significant difference associated with trenbolone dose.

Table E14T MHist-1. Descriptive statistics of the proportion of testes cells in each developmental stage for males from the EPA 14-day trenbolone assay and results of the Kruskal-Wallis Test (df = 2) comparing treatments.

|	Control (n = 8)			Low (n = 8)			High (n = 8)			Kruskal-Wallis	
Stage	**Mean**	**Stdev**	**CV**	**Mean**	**Stdev**	**CV**	**Value**	**Stdev**	**CV**	**H**	**p**
1	0	0	--	0	0	--	0	0	--	–	–
2a	0.006	0.005	72%	0.003	0.004	129%	0.003	0.003	113%	3.49	0.175
2b	0.018	0.008	41%	0.010	0.006	62%	0.013	0.012	92%	3.24	0.198
3a	0.167	0.060	36%	0.186	0.062	33%	0.045	0.037	80%	12.32	0.002**
3b	0.250	0.084	34%	0.204	0.095	47%	0.181	0.089	49%	2.22	0.329
4	0.173	0.056	33%	0.156	0.044	29%	0.127	0.111	88%	2.27	0.322
5	0.386	0.178	46%	0.441	0.184	42%	0.632	0.237	38%	4.75	0.093
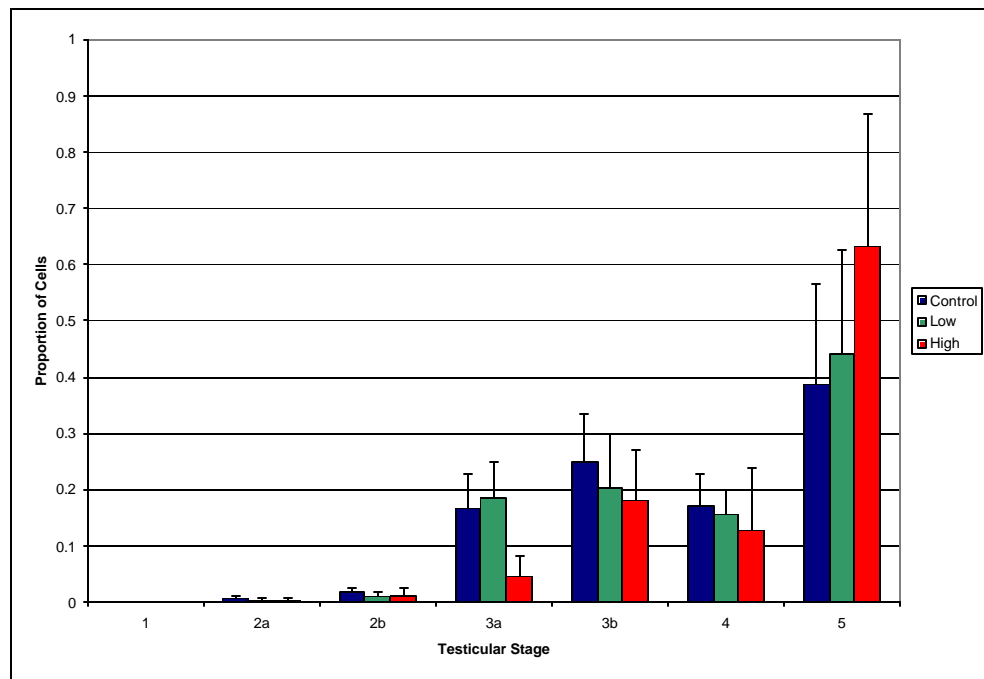
Figure E14T MHist-1. Frequency histogram showing the quantitative developmental staging of testes for each treatment of the EPA 14-day trenbolone assay. For each treatment, the columns represent the grand mean proportion of cells in each stage and the bars represent the standard deviation.

Tubule Diameter—The average diameter of the seminiferous tubules of males from the Control treatment ranged from 103.6 μm to 136.4 μm (**Figure E14T MHist-2**). Tubule diameters of males from the two test concentrations ranged from 90.8 μm to 172.8 μm. A significant difference in the mean tubule diameter per treatment (**Table E14T MHist-2**) was detected (Kruskal-Wallis, $H = 6.86$, $p = 0.032$, df = 2). However, the pattern of significance (Low < Control < High) was not consistent with the trenbolone dose. The achieved power for

this endpoint was 24%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 33 (**Table E14T MHist-2**).

Table E14T MHist-2. Summary statistics, detectable difference, and power estimates for male seminiferous tubule diameter data for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	125.4	13.6	11%	24%	33
low	8	112.1	20.7	18%		
high	8	138.6	23.4	17%		

[1] Calculated from natural log transformed data; with sample size = 2.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
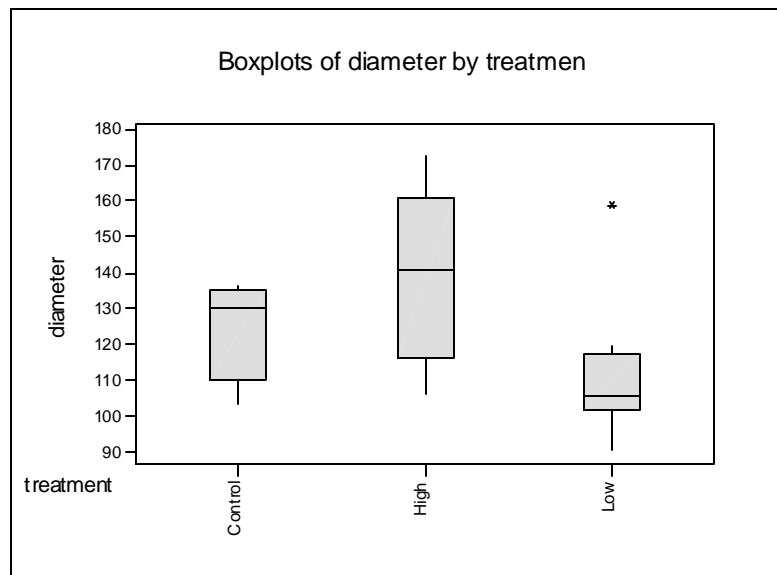

Figure E14T MHist-2. Boxplot of male seminiferous tubule diameter (mm) by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

Observations—No interstitial Sertoli cell proliferation and no Leydig cell proliferation was observed for any treatment. No testicular atrophy was recorded and no ovatestes were observed for any treatment.

Vitellogenin
Vitellogenin concentrations in Control treatment females used during the EPA 14-Day trenbolone assay ranged from 1,385,500 ng/mL to 3,722,500 ng/mL (**Figure E14T VTG-1**). Among females exposed to the two trenbolone concentrations, vitellogenin concentrations ranged from 0 ng/mL (not detected) to 5,847,500 ng/mL. Significant differences in the mean vitellogenin concentration per treatment (**Table E14T VTG-1**) were detected (Kruskal-Wallis, *H*

= 31.72, p = <0.001, df = 2). Vitellogenin concentrations in Control-treatment females and low-concentration females were significantly greater than those in females exposed to the High trenbolone concentrations. The achieved power for this endpoint was 100%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 4 (**Table E14T VTG-1**).

Table E14T VTG-1. Summary statistics, detectable difference, and power estimates for female vitellogenin concentrations (ng/mL) for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	16	2,608,969	592,805	23%	100%	4
low	14	2,237,993	1,798,862	80%		
high	16	40,910	79,868	195%		

[1] Calculated from natural log transformed data; with sample size = 14.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
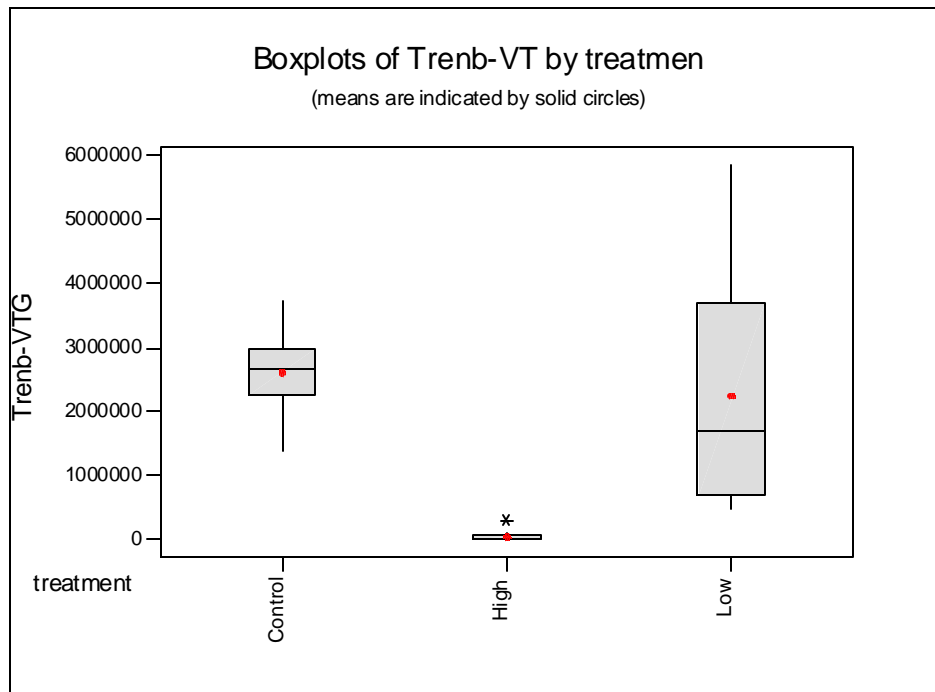

Figure E14T VTG-1. Boxplot of female vitellogenin concentration (ng/mL) by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, the asterisk represents a probable outlier.

Vitellogenin concentrations in Control treatment males used during the EPA 14-Day trenbolone assay ranged from 0 ng/mL (not detected) to 8,813 ng/mL (**Figure E14T VTG-1**). Among males exposed to the two trenbolone concentrations, vitellogenin concentrations ranged

from 0 ng/mL to 2,333 ng/mL. No significant differences in the mean vitellogenin concentration per treatment (**Table E14T VTG-1**) were detected (Kruskal-Wallis, $H = 1.26$, $p = 0.532$, df = 2). The achieved power for this endpoint was 12%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 77 (**Table E14T VTG-1**).

Table E14T VTG-1. Summary statistics, detectable difference, and power estimates for male vitellogenin concentrations (ng/mL) for the EPA 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	1,811	3,061	169%	14%	66
low	8	774	1,033	133%		
high	8	447	226	51%		

[1] Calculated from natural log transformed data; with sample size = 8.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
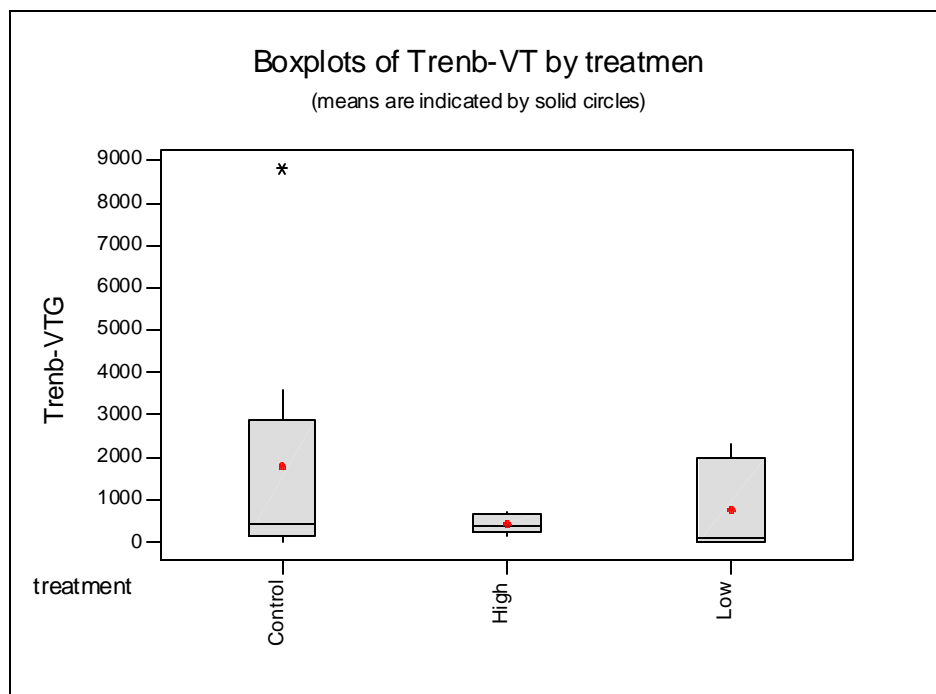

Figure E14T VTG-1. Boxplot of male vitellogenin concentration (ng/mL) by treatment for the EPA 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, the asterisk represents a probable outlier.

Plasma Steroid Concentrations

3.2.4 Non-spawning Adult 14-day Assay for trenbolone

The Non-spawning Adult 14-day trenbolone assay was conducted from February 10, 2003 to February 24, 2003.

Chemistry

Trenbolone concentrations were measured in each replicate on Day 0, 7, and 14 for the all concentrations and also on Day 3 for the Low concentration (**Figure OECD14M Chem-1**) during the Non-spawning Adult 14-D trenbolone exposure. The mean percent difference of the Low, Medium, and High concentrations achieved during the test versus the target concentrations were 29% , 11%, and 15%, respectively (**Table OECD14M Chem-1**).

Table OECD14T Chem-1. Summary statistics for trenbolone concentrations (ng/L) during the Non-spawning Adult 14-Day assay.

	Control	Low	Medium	High
Target	0.051 [1]	0.1	0.5	1
mean	0.051 U	0.071	0.445 [2]	0.86
stdev	0	0.026	0.045	0.11
min	0.051 U	0.038	0.390	0.73
max	0.051 U	0.109	0.510	1.02
median	0.051 U	0.059	0.430	0.86
Q1	0.051 U	0.054	0.423	0.77
Q3	0.051 U	0.091	0.475	0.91
Mean % Δ	–	29%	11%	15%

[1] Equal to the MDL.
[2] Concentrations in some replicates were estimated at levels less than the MDL.
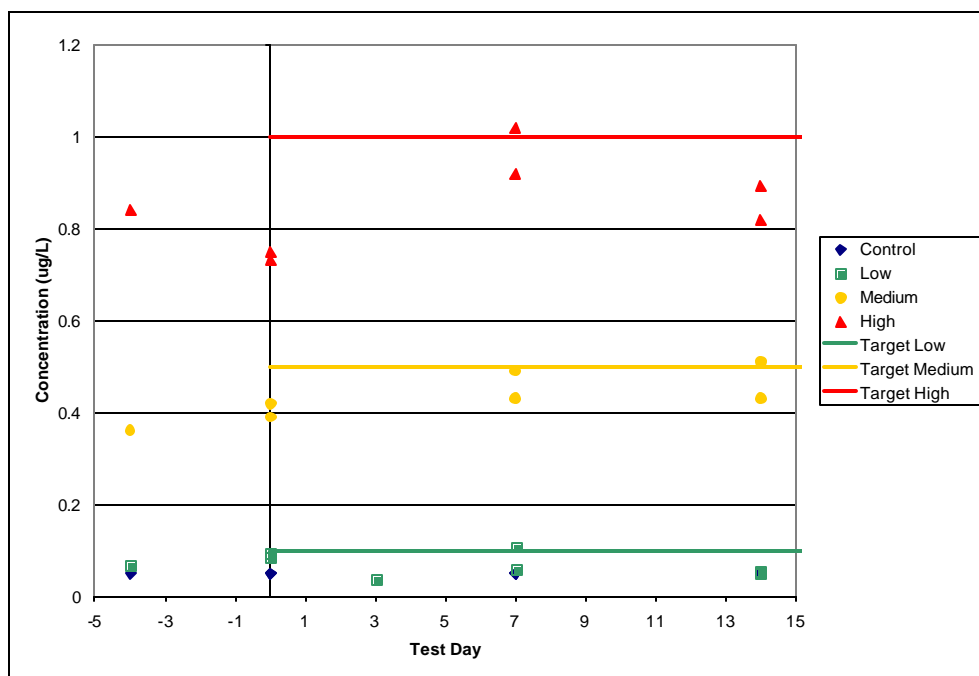U Not detected at concentrations greater than the MDL.

Figure OECD14T Chem-1. trenbolone concentrations measured in the High and Low-concentration treatments versus the target concentrations (1.0 ng/L and 0.1 ng/L) during the Non-spawning Adult 14-day trenbolone assay.

Test Conditions Summary

       All water quality, water chemistry, and light intensity parameters were within the acceptable ranges during the Non-spawning Adult 14-day trenbolone assay (Table OECDT WQ-1). Flow rates were monitored regularly, and were adjusted if they were observed to be out of the target range (see Methods).

Table OECDT WQ-1. Summary of water quality, water chemistry, flow rate, and light intensity data for the Non-spawning Adult 14-day trenbolone assay.

	Temperature (°C)		Dissolved Oxygen (mg/L)		pH		Hardness (mg/L)		Alkalinity (mg/L)		Ammonia (mg/L)		Light Intensity (lux)			
	min	max	min	max	min	max	min	max	min	max	min	max	min	max	mean	SD
Target Range	24	26	>4.9		6.5	9.0	>140		>20		<0.5		540	1080		
Control	25	25	5.7	7.6	7.4	8.0	140	204	150	175	0.1	0.4	952	968	960	11
Low	24	25	6.0	5.7	7.2	8.1	170	221	150	185	0.1	0.3	919	973	946	38
Medium	25	26	5.8	7.8	7.5	8.2	170	204	150	165	0.1	0.3	724	989	957	187
High	25	25	6.0	7.9	7.6	8.2	170	204	150	175	0.1	0.2	916	934	920	11

Survival

       Total survival in the Control and all three test concentrations during the Non-spawning Adult 14-day trenbolone assay was 100%.

Body Weight and Length

The body weight of females used in the Non-spawning Adult 14-day trenbolone assay ranged from 1.4 g to 3.3 g (**Figure OECD14T BWL-1**). There were no significant differences in mean body weight (natural log transformed) among treatments (Kruskal-Wallis, $H = 4.05$, $p = 0.256$, df = 3). The achieved power for this endpoint was 35%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 26 (**Table OECDT BWL-1**).

The body (fork) length of females used in the Non-spawning Adult 14-day trenbolone assay ranged from 47 mm to 63 mm (**Figure OECD14T BWL-2**). There were no significant differences in mean body weight (natural log transformed) among treatments (Kruskal-Wallis, $H = 5.90$, $p = 0.117$, df = 3). The achieved power for this endpoint was 33%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 28 (**Table OECDT BWL-2**).

Table OECDT BWL-1. Summary statistics, detectable difference, and power estimates for female body weight (g) data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	1.9	0.3	16%	35%	26
low	10	2.1	0.4	19%		
medium	10	2.3	0.5	22%		
high	10	2.1	0.5	23%		

[1] Calculated from natural log transformed data; with sample size = 10.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.

Boxplots of fish_wgt by treatmen
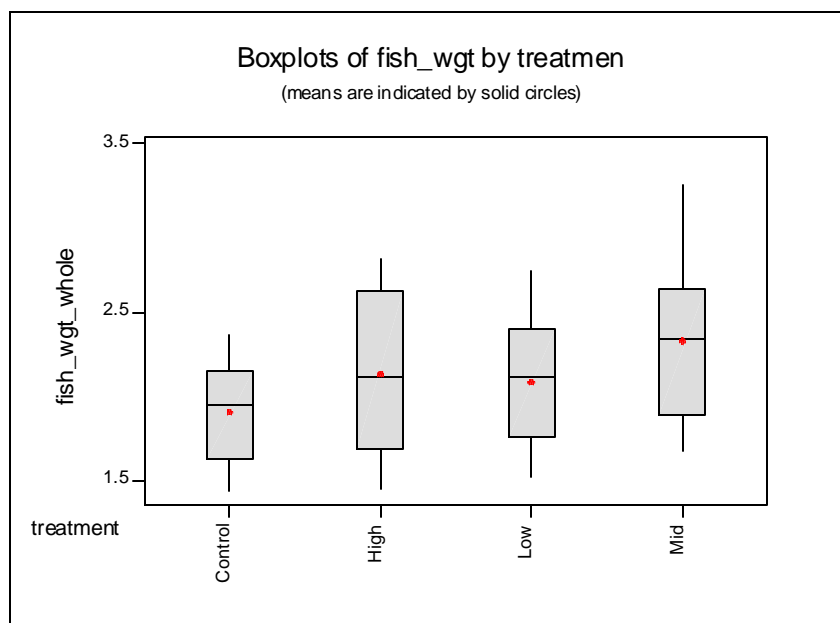(means are indicated by solid circles)

Figure OECD14T BWL-1. Boxplot of female body weight (g) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Table OECDT BWL-2. Summary statistics, detectable difference, and power estimates for female body length (mm) data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	51.8	3.9	7%	33%	28
low	10	54.0	3.1	6%		
medium	10	55.2	3.3	6%		
high	10	52.8	4.9	9%		

[1] Calculated from natural log transformed data; with sample size = 10.

[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
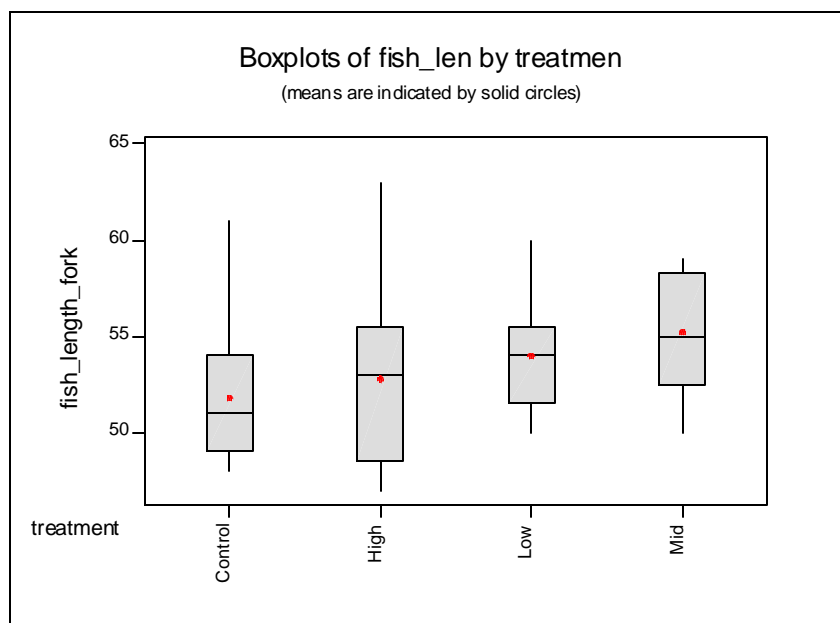
Figure OECD14T BWL-2. Boxplot of female body length (mm) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

The body weight of males used in the EPA 14-day trenbolone assay ranged from 2.4 g to 6.5 g (**Figure OECD14T BWL-3**). There were no significant differences in mean body weight among treatments (Kruskal-Wallis, $H = 5.34$, $p = 0.148$, df = 3). The achieved power for this endpoint was 48%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 19 (**Table OECDT BWL-3**).

The body (fork) length of males used in the Non-spawning Adult 14-day trenbolone assay ranged from 51 mm to 79 mm (**Figure OECD14T BWL-4**). There were no significant differences in mean body weight (natural log transformed) among treatments (Kruskal-Wallis, $H = 6.72$, $p = 0.081$, df = 3). The achieved power for this endpoint was 48%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 19 (**Table OECDT BWL-4**).

Table OECDT BWL-3. Summary statistics, detectable difference, and power estimates for male body weight (g) data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	3.7	0.8	21%	48%	19
low	10	4.0	0.8	21%		
medium	10	4.7	1.1	23%		
high	10	4.4	0.9	21%		

[1] Calculated from natural log transformed data; with sample size = 10.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.

Boxplots of fish_wgt by treatmen
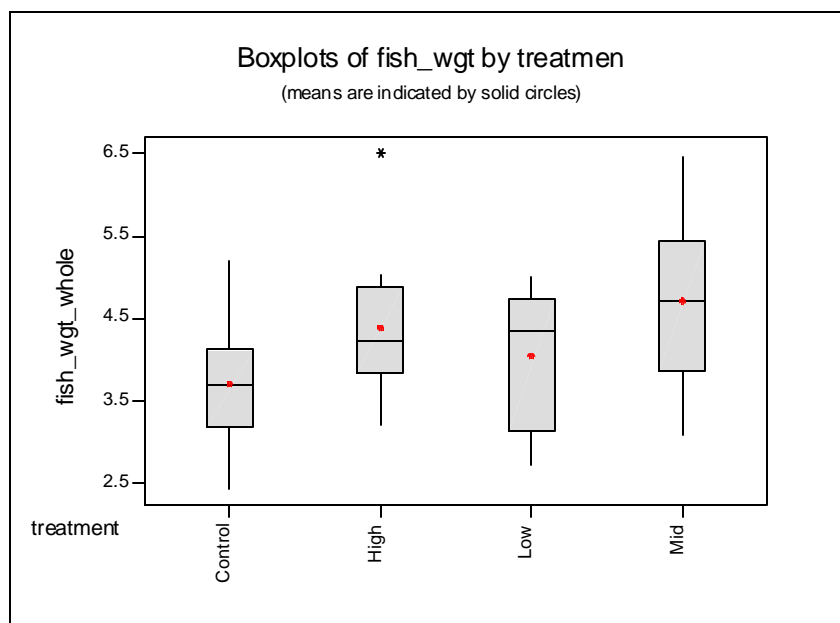(means are indicated by solid circles)

Figure OECD14T BWL-3. Boxplot of male body weight (g) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

Table OECDT BWL-4. Summary statistics, detectable difference, and power estimates for male body length (mm) data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	59.9	5.4	9%	48%	19
low	10	65.4	4.8	7%		
medium	10	65.8	7.1	11%		
high	10	64.8	3.2	5%		

[1] Calculated from natural log transformed data; with sample size = 10.

[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
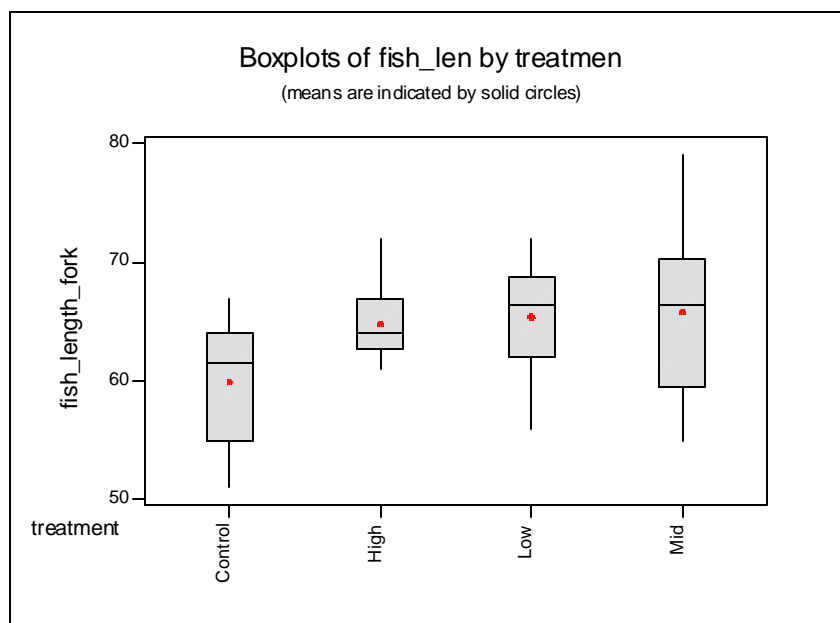
Figure OECD14T BWL-4. Boxplot of male body length (mm) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Gonadosomatic Index

The range of GSI values calculated for females in the all treatments varied from two- to ten-fold (**Figure OECD14T GSI-1**), and the overall within-treatment variability was moderate (CVs = 22% –38%; **Table OECD14T GSI-1**). The highest female GSI value was 24.9 (one fish in the High concentration), but several fish had GSI values >20. There were no significant differences in mean GSI values (**Table OECD14T GSI-1**) among treatments (Kruskal-Wallis, $H = 3.05$, $p = 0.384$, df = 3). The achieved power for this endpoint was 18%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 55 (**Table OECD14T GSI-1**).

Table OECDT GSI-1. Summary statistics, detectable difference, and power estimates for female gonadosomatic index data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	13.2	5.0	38%	18%	55
low	10	13.8	3.5	25%		
medium	10	16.4	3.6	22%		
high	10	14.8	7.4	50%		

[1] Calculated from arcsine square-root transformed data; with sample size = 10.

[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
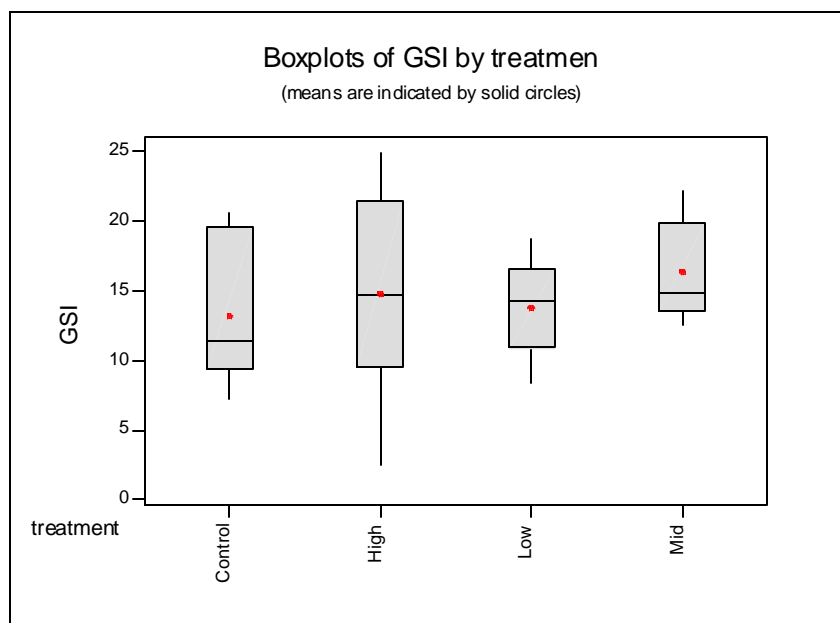
Figure OECD14T GSI-1. Boxplot of female GSI by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

The range of most GSI values calculated for males during the Non-spawning Adult 14-day trenbolone assay, varying from 0.8 to 2.0 (**Figure OECD14T GSI-2**), which approximates the typical range for reproductively-active male fathead minnows. Within-treatment variability was moderate, as indicated by CVs that ranged from 24% to 59%; **Table OECD14T GSI-2**). The highest and lowest male GSI values were 2.2–2.3 (for one fish each in the Low and High concentrations) and 0.2–0.3 (one fish each in the Low and Medium concentrations), respectively. However, there were no significant differences among treatments in mean GSI values (**Table OECD14T GSI-2**) among treatments (Kruskal-Wallis, $H = 4.60$, $p = 0.203$, df = 3). The achieved power for this endpoint was 16%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 64 (**Table E14T GSI-2**).

Table OECD14T GSI-2. Summary statistics, detectable difference, and power estimates for male gonadosomatic index data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	1.21	0.41	34%	16%	64
low	10	1.11	0.66	59%		
medium	10	1.34	0.56	42%		
high	10	1.60	0.38	24%		

[1] Calculated from arcsine square-root transformed data; with sample size = 10.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on arcsine square-root transformed data.
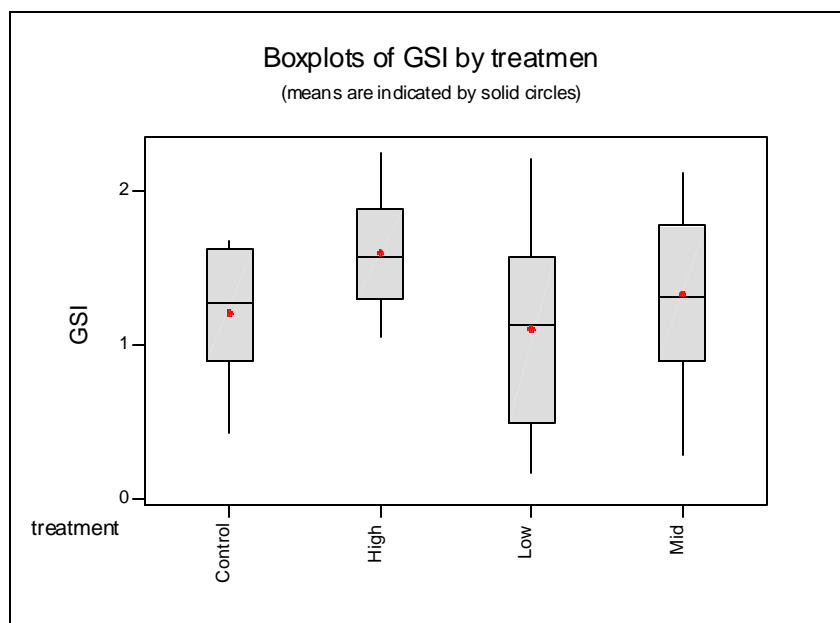
Figure OECD14T GSI-2. Boxplot of male GSI by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, and the circle is the mean value.

Female Gonad Histology

General Ovary Staging—Statistical analysis of the mean ovarian staging from 12 microscopic fields per female in the Non-spawning Adult 14-day trenbolone assay revealed a significant difference among treatments (Kruskal-Wallis, $H = 9.13$, $p = 0.028$, df = 2). The mean ovarian stage of females in the Control treatment was greater than that of females in the other three treatments.

Quantitative Ovarian Staging—One hundred cells in each of three sections per female were examined to quantitatively determine the developmental stage of the ovaries. Ova from fish in the Control treatment, Medium concentration, and High concentration ranged from Stage 1a to Stage 5 (see Methods for a description of the stages), whereas ova from females from the Low-concentration treatment showed Stage 1a to Stage 4 development (**Figure OECD14T FHist-1**). Variability within treatments for each stage was very high as indicated by CVs that ranged as high as 316% (**Table OECD14T FHist-1**). Although statistical analyses showed that there was a significant difference among treatments in the proportion of cells in developmental Stages 3 and 5, there were no significant differences among treatments in the proportion of cells in the developmental Stages 1a, 1b, 2, and 4 (**Table OECD14T FHist-1**). Therefore, there was no consistent pattern of significant difference associated with trenbolone dose.

Table OECD14T FHist-1. Descriptive statistics of the proportion of ovarian cells in each developmental stage for females from the Non-spawning Adult 14-day trenbolone assay and results of the Kruskal-Wallis Test (df = 2) comparing treatments.

| Stage	Control (n = 10)			Low (n = 10)			Medium (n = 10)			High (n = 10)			Kruskal-Wallis	
	Mean	Stdev	CV	Mean	Stdev	CV	Mean	Stdev	CV	Mean	Stdev	CV	*H*	*p*
1a	0.107	0.034	32%	0.086	0.029	33%	0.082	0.041	51%	0.069	0.047	68%	5.65	0.130
1b	0.370	0.085	23%	0.325	0.080	25%	0.257	0.080	31%	0.291	0.128	44%	7.70	0.053
2	0.135	0.048	35%	0.137	0.046	33%	0.108	0.043	40%	0.151	0.052	34%	3.31	0.346
3	0.096	0.047	48%	0.202	0.093	46%	0.236	0.072	31%	0.238	0.088	37%	15.03	0.002**
4	0.211	0.154	73%	0.233	0.116	50%	0.251	0.131	52%	0.197	0.134	68%	1.41	0.704
5	0.034	0.052	151%	0.000	0.000	–	0.002	0.006	316%	0.002	0.004	179%	9.28	0.026*
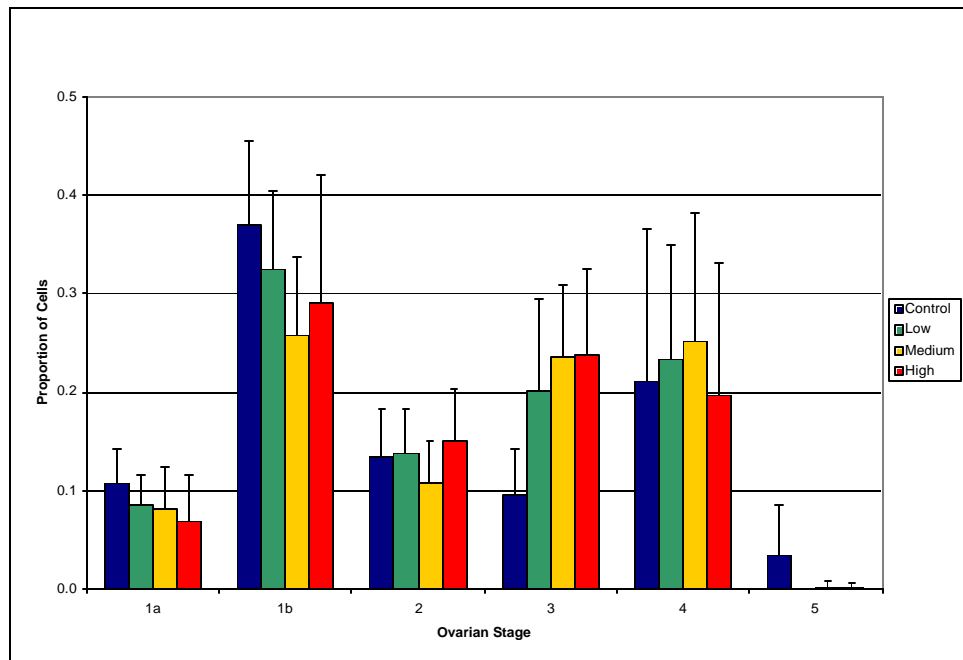
Figure OECD14T FHist-1. Frequency histogram showing the quantitative developmental staging of ovaries for each treatment of the Non-spawning Adult 14-day trenbolone assay. For each treatment, the columns represent the grand mean proportion of cells in each stage and the bars represent the standard deviation.

Atretic Follicles—The mean proportion of atretic follicles per 300 follicles (counted per fish) ranged from 0.01 follicles for females in the Low concentration to 0.05 follicles for females in the High and Medium concentrations (**Figure OECD21T FHist-2**). There were significant differences in the proportions of atretic follicles among treatments (Kruskal-Wallis, $H = 10.32$, $p = 0.016$, df = 2). The value for females in the Low concentration was lower than those for females in the other three treatments. However, there was no consistent pattern of significant difference associated with trenbolone dose.

Boxplots of atretic_follicles by treatmen

Figure OECD14T FHist-2. Boxplot of the proportion of atretic follicles per 300 follicles by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

Corpora Lutea—The mean proportion of corpora lutea per 300 follicles (counted per fish) ranged from 0.001 for females in the High concentration to 0.011 for females in the Control treatment (**Figure OECD14T FHist-3**). There were significant differences in the proportions of corpora lutea among treatments (Kruskal-Wallis, $H = 10.16$, $p = 0.017$, df = 2). The mean value for fish from the High concentration was less than that for the Low concentration, but not than that for the Control treatment.
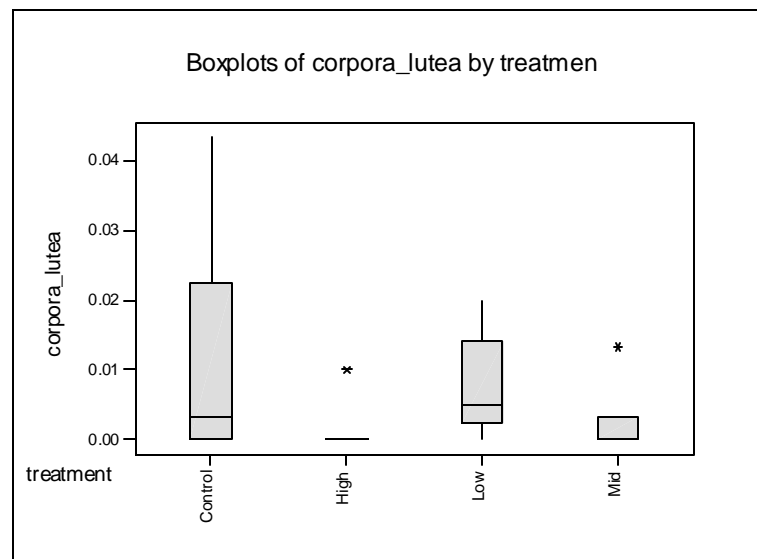

Boxplots of corpora_lutea by treatmen

Figure OECD14T FHist-3. Boxplot of the proportion of corpora lutea per 300 follicles by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the me an value, and asterisks represent probable outliers.

Male Gonad Histology

Testes Staging by Microscopic Field —Testes from males exposed to trenbolone during the Non-spawning Adult 14-day trenbolone assay were examined to determine the general developmental condition. Males in all treatments had well-developed testes with most showing Stage 4 and Stage 5 development (see Methods for description of developmental stages). All of the 120 microscopic fields examined in the 10 Control treatment males showed Stage 4 (58 fields) or Stage 5 (62 fields) development. All of the 108 microscopic fields examined in the 9 Low-concentration treatment males showed Stage 4 (74 fields) or Stage 5 (34 fields) development. All of the 120 microscopic fields examined in the 10 Medium concentration treatment males showed Stage 4 (50 fields) or Stage 5 (70 fields) development. All of the 120 microscopic fields examined in the 10 High-concentration treatment males showed Stage 4 (76 fields) or Stage 5 (44 fields) development. Statistical analysis of the mean staging from 12 microscopic fields per fish revealed no significant differences among treatments (Kruskal-Wallis, $H = 4.24$, $p = 0.236$, df = 2).

Quantitative Testicular Staging—One hundred cells in each of three sections per male were examined to quantitatively determine the developmental condition of the testes. The developmental stage of the testes from fish in all treatments, except the Low concentration, ranged from Stage 2a to Stage 5 (**Figure OECD 14T MHist-1**). The developmental stage of the testes from fish in the Low concentration ranged from Stage 2b to Stage 5. Variability within treatments for each stage was very high as indicated by CVs that ranged as high as 211% (**Table OECD 14T MHist-1**). There were no significant differences among treatments in the proportion of cells in developmental Stages 2a, 2b, 3a, 3b, 4, and 5 (**Table OECD 14T MHist-1**). Therefore, there was no consistent pattern of significant difference associated with trenbolone dose.

Table OECD 14T MHist-1. Descriptive statistics of the proportion of testes cells in each developmental stage for males from the Non-spawning Adult 14-day trenbolone assay and results of the Kruskal-Wallis Test (df = 2) comparing treatments.

| Stage | Control (n = 10) | | | Low (n = 9) | | | Medium (n = 10) | | | High (n = 10) | | | Kruskal-Wallis | |
	Mean	Stdev	CV	Mean	Stdev	CV	Mean	Stdev	CV	Mean	Stdev	CV	H	p
1	0	0	–	0	0	–	0	0	–	0	0	–	–	–
2a	0.001	0.002	175%	0	0	–	0.002	0.003	194%	0.001	0.001	211%	3.44	0.329
2b	0.009	0.013	150%	0.010	0.012	121%	0.005	0.008	159%	0.007	0.008	125%	1.10	0.777
3a	0.027	0.032	119%	0.066	0.055	84%	0.046	0.076	165%	0.055	0.031	56%	5.69	0.128
3b	0.116	0.111	95%	0.149	0.082	55%	0.152	0.162	107%	0.117	0.093	79%	1.62	0.654
4	0.082	0.084	102%	0.090	0.061	67%	0.064	0.074	115%	0.087	0.066	76%	1.77	0.621
5	0.765	0.224	29%	0.685	0.166	24%	0.731	0.305	42%	0.700	0.170	24%	1.79	0.618
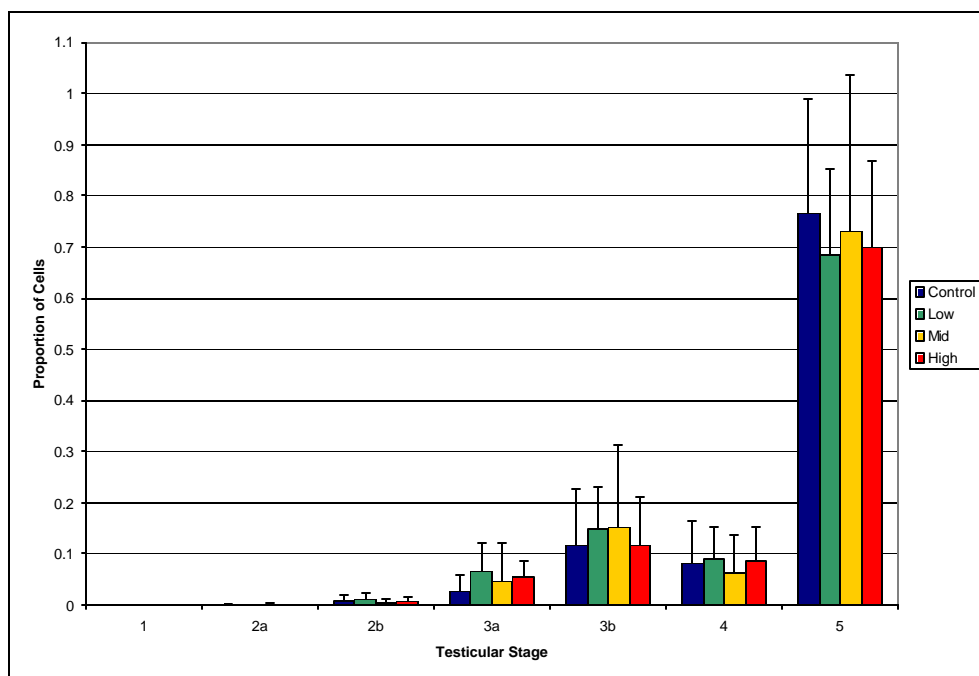
Figure OECD 14T MHist-1. Frequency histogram showing the quantitative developmental staging of testes for each treatment of the Non-spawning Adult 14-day trenbolone assay. For each treatment, the columns represent the grand mean proportion of cells in each stage and the bars represent the standard deviation.

Tubule Diameter—The average diameter of the seminiferous tubules of males from the Control treatment ranged from 85.3 µm to 236.7 µm (**Figure OECD 14T MHist-2**). Tubule diameters of males from the three test concentrations ranged from 87.2 µm to 250.0 µm. No significant differences in the mean tubule diameter per treatment (**Table OECD 14T MHist-2**) were detected (Kruskal-Wallis, $H = 2.54$, $p = 0.468$, df = 2). The achieved power for this endpoint was 13%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 78 (**Table OECD 14T MHist-2**).

Table OECD14T MHist-2. Summary statistics, detectable difference, and power estimates for male seminiferous tubule diameter data for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	10	159.3	54.8	10	13%	78
low	9	133.3	41.8	9		
medium	10	162.2	42.4	10		
high	10	160.3	39.5	10		

[1] Calculated from natural log transformed data; with sample size = 9.

[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
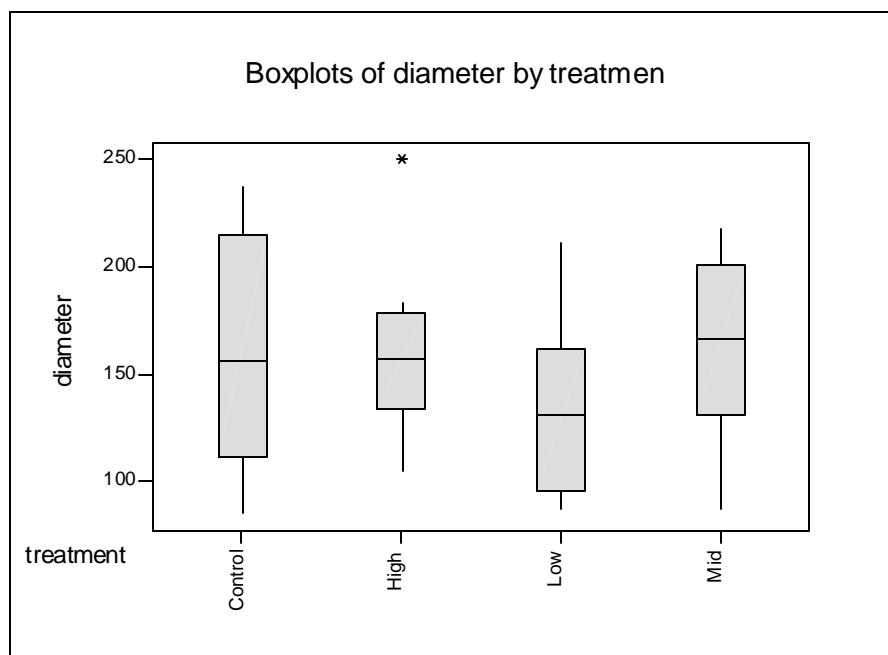
Figure OECD14T MHist-2. Boxplot of male seminiferous tubule diameter (mm) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, and the asterisk represents a probable outlier.

Observations—Two males in the Medium concentration treatment showed disorganized tubular components. No interstitial Sertoli cell proliferation and no Leydig cell proliferation was observed for any treatment. No testicular atrophy was recorded and no ovatestes were observed for any treatment.

Vitellogenin
Vitellogenin concentrations in Control treatment females used during the Non-spawning Adult 14-Day trenbolone assay ranged from 1,409,500 ng/mL to 4,440,000 ng/mL (**Figure OECD14T VTG-1**). Among most females exposed to the three trenbolone concentrations, vitellogenin concentrations ranged from 0 ng/mL (not detected) to 915,750 ng/mL. Two females exposed to the low trenbolone concentration had vitellogenin concentrations of 1,549,500 ng/mL and 3,821,000 ng/mL. Significant differences in the mean vitellogenin concentration among treatments (**Table OECD14T VTG-1**) were detected (Kruskal-Wallis, $H = 23.23$, $p = <0.001$, df $= 2$). Mean vitellogenin concentrations among Control treatment females were significantly greater than the mean concentrations for females exposed to the all three trenbolone concentrations. The achieved power for this endpoint was 98%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 5 (**Table OECD14T VTG-1**).

Table OECD14T VTG-1. Summary statistics, detectable difference, and power estimates for female vitellogenin concentrations (ng/mL) for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	8	2,810,125	1,019,160	36%	96%	6
low	10	907,905	1,123,918	124%		
medium	9	125,801	241,178	192%		
high	9	112,255	145,135	129%		

[1] Calculated from natural log transformed data; with sample size = 8.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.

Boxplots of Trenb-VT by treatmen
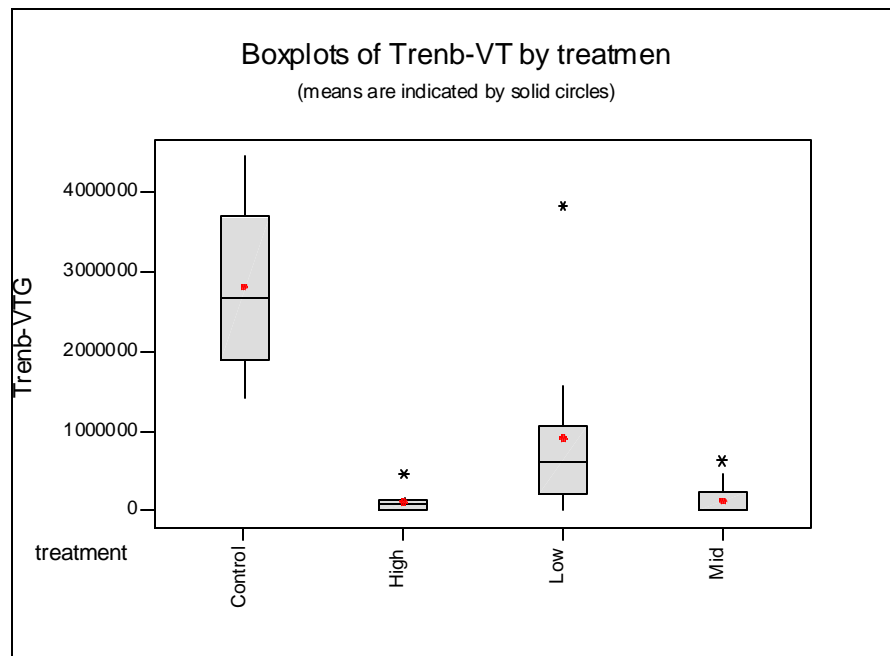(means are indicated by solid circles)

Figure OECD14T VTG-1. Boxplot of female vitellogenin concentration (ng/mL) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, asterisks represent probable outliers.

Vitellogenin concentrations in Control-treatment males used during the Non-spawning Adult 14-Day trenbolone assay ranged from 0 ng/mL (not detected) to 7,672 ng/mL (**Figure OECD14T VTG-1**). Vitellogenin was not detected above the MDL for four of the eight males in the control treatment. Among most males exposed to the three trenbolone concentrations, vitellogenin concentrations ranged from 0 ng/mL (not detected) to 1,020 ng/mL. Vitellogenin concentrations were not detected above the MDL for 18 of the 20 males exposed to the low and medium trenbolone concentrations. One male exposed to the High trenbolone concentration had a vitellogenin concentration of 11,390 ng/mL. Significant differences in the mean vitellogenin

concentration per treatment (**Table OECD14T VTG-1**) were detected (Kruskal-Wallis, $H =$ 16.81, $p = 0.001$, df = 2). Vitellogenin concentrations in males form the Control treatment and the High concentration ranked higher than those from males in the Low and Medium concentrations. The achieved power for this endpoint was 41%, and the sample size required to detect a significant difference from the Control treatment at 80% power was 20 (**Table OECD14T VTG-1**).

Table OECD14T VTG-1. Summary statistics, detectable difference, and power estimates for male vitellogenin concentrations (ng/mL) for the Non-spawning Adult 14-day trenbolone assay.

Level	N	Mean	Stdev	CV	Achieved Power [1]	Sample Size Required [2]
control	9	1,337	2,561	192%	41%	20
low	10	102	323	317%		
medium	10	30	96	320%		
high	10	1,449	3,500	242%		

[1] Calculated from natural log transformed data; with sample size = 9.
[2] To detect a significant difference from control treatment based on maximum achieved absolute difference; calculated on natural log transformed data.
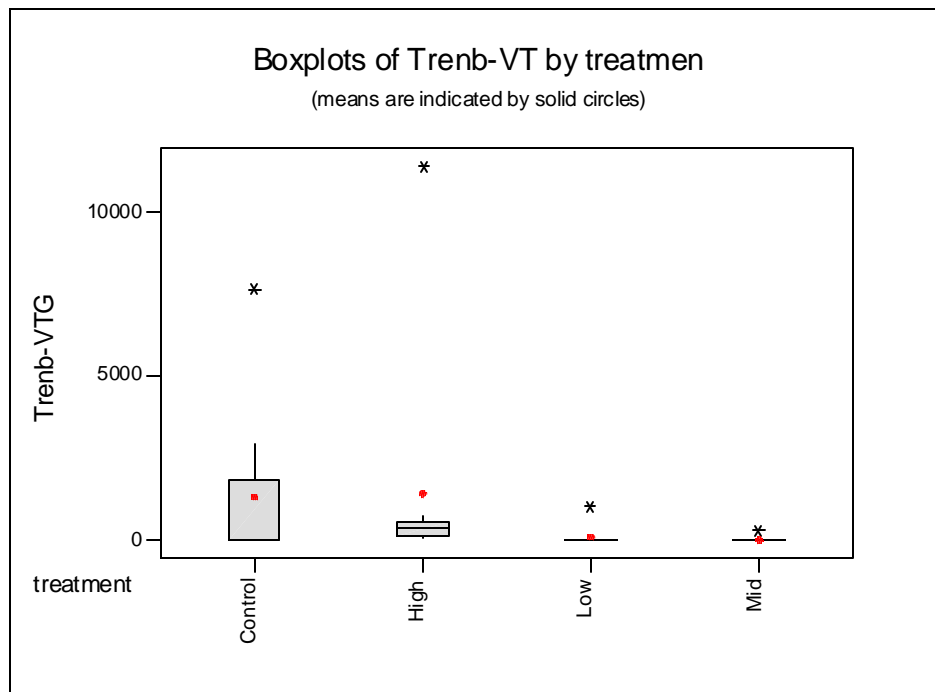

Figure OECD14T VTG-1. Boxplot of male vitellogenin concentration (ng/mL) by treatment for the Non-spawning Adult 14-day trenbolone assay. The box represents the interquartile range, whiskers represent the data range, the horizontal line is the median value, the circle is the mean value, asterisks represent probable outliers.

Plasma Steroid Concentrations